Radiation reaction techniques in general relativity

by Jordan Emrys Moxon
August 2018

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

Radiation reaction techniques in general relativity

Jordan Emrys Moxon, Ph.D.

Cornell University 2018

Abstract

This dissertation presents several results which pertain to self-interaction effects in general relativity. I first present a detailed review of the physics of gravitational radiation, and compact binaries in general, which provide the key motivation for detailed exploration of the physical processes and modelling of radiation reaction effects. In particular, extreme mass ratio inspirals (EMRIs) and intermediate mass ratio inspirals (IMRIs) are a promising source for current and near-future gravitational wave detectors, and detailed modeling of the resulting waveforms requires a deep understanding of the physics of self interactions.

The first project which I present in this dissertation is the first derivation of the second-order self force for scalar and electromagnetic fields in flat spacetime. In the process of developing this derivation, several important mathematical subtleties in the definitions of bulk body parameters emerged, which become important at the same order at which gravitational self force computations are considered. Additionally, this self force is of intrinsic interest for charges accelerated by high-powered lasers and in astrophysical systems. Our derivation lays the ground work for future derivations of high order self force in curved spacetime, which may be important for testing alternative theories of gravity.

The second project, which comprises the majority of this dissertation, details the development of a tapestry of approximations for highly accurate simulation of high mass ratio inspirals. Due to the precision requirements for waveform templates important for LISA data analysis, one of the significant goals of self force calculations is to compute waveforms that track the phase of the long evolution of EMRIs to a precision far better than one radian. One of the methods is a multiscale expansion which exploits the separation of scales between the slow radiation-reaction time and the fast orbital time. This dissertation discusses in detail the mathematical techniques of the multiscale approximation method and other approximations in the various regions of the spacetime. The techniques, which I develop in collaboration with Éanna Flanagan, Tanja Hin-

derer, and Adam Pound, comprise the only currently available method of assuring sub-radian accuracy in the computed waveform.

The last project presented in this dissertation develops new techniques for quantizing theories which exhibit gauge degrees of freedom. We suggest a modification of the well-known Dirac bracket formalism. The alternative formalism may be more useful than the more frequently used methods of BRST quantization for derivations of local effects of dressed states. In addition, the modified Dirac bracket may be more computationally simple than the original, which may be of use for computations in intricate theories.

Biographical sketch

Jordan Moxon was born and spent his childhood near the small town of Rhinelander in Northern Wisconsin. He matriculated to the University of Wisconsin - Madison in 2008, where he enjoyed the ability to rapidly accrete majors. By the conclusion of his undergraduate education, he had completed a triple major in Physics, Computer Science, and Mathematics. During his time at the University of Wisconsin, he had multiple roles with several research groups within the University. At various points, he acted as the student systems administrator for the servers of the Phillip's lab in Biochemistry, and as an undergraduate research assistant for the Gleicher lab in Computer science and for the McDermott lab in Physics. His undergraduate physics studies culminated in a senior thesis studying inflationary perturbations under Prof. Gary Shiu.

Jordan is currently a Ph.D. candidate at Cornell University, where he conducts research on topics in gravitational physics with his Ph.D. advisor, Prof. Éanna Flanagan. Jordan's primary research focus is now analytical developments regarding the self-force problem, the goal of which is a better understanding of techniques for generating gravitational waveforms that arise from high mass ratio binary black hole inspirals. Jordan's research has also included investigations of quantization techniques in gauge theories and calculations of perturbative quantum gravity effects. In addition, Jordan has made significant contributions during his years as a Teaching Assistant, helping to update lab equipment, considering alterations to lab curricula, and acting as co-instructor to Prof. Seamus Davis during his final year of Ph.D. work.

Acknowledgements

My Ph.D. advisor, Prof. Éanna Flanagan: First, I would like to express my gratitude for the time and effort Prof. Éanna Flanagan has put forth during my graduate education. Early in my graduate career, he acted as a superb instructor, infusing both a deep mathematical understanding of relativistic effects and a keen understanding of scales and estimation techniques, which are key to identifying salient aspects of a system or calculation. Later, collaborating with him has been a joy, due to the fascinating topics of research, the challenge of computations, and the engaging discussions of results and calculational details.

My Collaborators, Adam Pound and Tanja Hinderer: I also wish to thank my long-time collaborators, Adam Pound and Tanja Hinderer, with whom the research on multiscale approximation methods was conducted. Working with them has been extremely valuable and instructive. The quality and extent of the multiscale approximation results which I discuss in this dissertation are the fruits of the combined effort, and without them, the scope of the project would not have been possible.

My family and friends: I thank all of my friends and family for their continued support over the past several years of my graduate education. Particularly, I want to thank Jennifer Chu for her enduring assistance and encouragement. Additionally, I thank my parents, John and Mary Moxon for their years of support and their persistant enthusiasm in listening to even the most arcane explanations of relativistic research progress. I thank all of the fellow research group members I have had over the years, including Barry Wardell, Peter Taylor, Ulysses Machado, Leo Stein, David Nichols, Jolyon Bloomfield, Justin Vines, Alex Grant, Kartik Prabhu, and Ibrahim Shehzad, for productive discussions and exposure to interesting research. I also thank all of my close friends in graduate school for many nights of board games and late-night arguments. I thank Neal Reynolds for embarking on a crazy side-project with me, which kept us both sane.

Funding: This work was supported in part by NSF grants PHY-1404105 and PHY-1707800

Contents

Chapter 1

A review of gravitational wave theory in general relativity

1.1 Gravitational waves and searches

General relativity predicts that spacetime is a nonlinear, dynamic field, and the field equations of general relativity support propagating wavelike solutions. Gravitational waves were first predicted as a consequence of the equation of motion for general relativity over a century ago by Einstein [1]. The prediction of propagating gravitational perturbations is a core implication of general relativity, and the recent detection of gravitational waves by the LIGO detectors [2] offer powerful insights both into the fundamental properties of gravitation and the nature of the sources which generated the gravitational disturbances.

The search for Gravitational waves has a long history. When gravitational waves were first proposed, it was a subject of great debate whether they could to transfer energy to matter. In the 1957 Chapel Hill conference [3], it was cogently argued by Richard Feynmann via a 'sticky bead' thought experiment, that indeed gravitational waves should carry energy and at least in principle, a device could be constructed to detect gravitational waves. It is notable, though, that the strength of the effect of gravitational waves is so weak that Einstein's thoughts were that "it is obvious that [The energy radiated by a system per unit time] has, in all imaginable cases, a practically vanishing value" (from a translation [4]).

When a set of observers is permitted to move freely, the effect of a gravitational wave passage

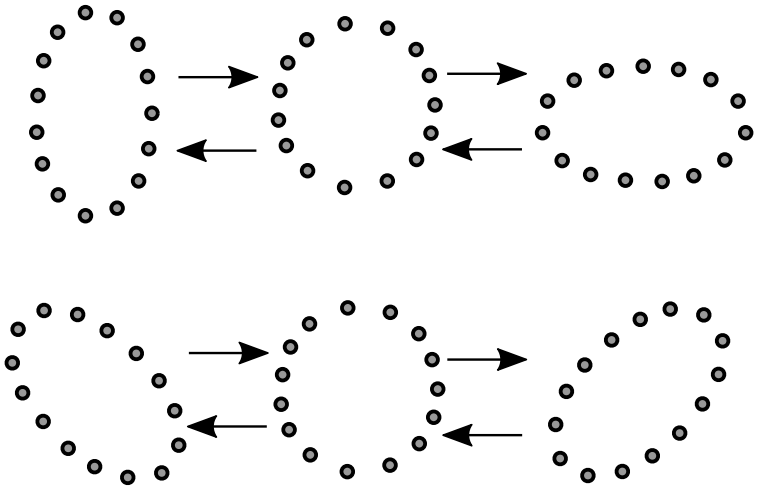

Figure 1.1: A visualization of the detectable consequence of a gravitational wave passage. The ring of observers will be periodically stretched and squeezed along the principal axes of the gravitational wave polarization. The polarization depicted in the top ring is referred to as the +("plus") mode, and the polarization depicted in the bottom ring is referred to as the ×("cross") mode

is a small transverse oscillation in the observed displacements. If arranged in a circular ring, the ring will distort to an ellipse along the axes of the gravitational wave polarization (see Fig. 1.1). If instead, however, a set of objects are bound together as a large resonant oscillator, the gravitational wave will give rise to a force stretching and squeezing the group along orthogonal axes. In fact, an early experiment to attempt to detect gravitational waves involved an approximately four metric ton aluminum bar, which was initially claimed to have detected gravitational waves by virtue of its resonant properties [5]. However, the quality factor of the bar and the sensitivity of the readout was far lower than would be necessary to detect astrophyical sources, and it was discovered that the experimental procedure was not robust.

The first successful indication of gravitational radiation dissipation was found in the Hulse-Taylor pulsar [6]. The pulsar was found in a close binary orbit, in which significant orbital energy and angular momentum is lost to gravitational radiation. As a result, general relativity provided a specific prediction for the rate of the orbital evolution. The resulting orbital evolution was then readily observable by the predictable radio pulses from the pulsar system. The extremely tight agreement (see Fig. 1.2) with the predictions of general relativity were aptly hailed as a triumph both of observational astronomy and of the accuracy of general relativity for describing astrophysical

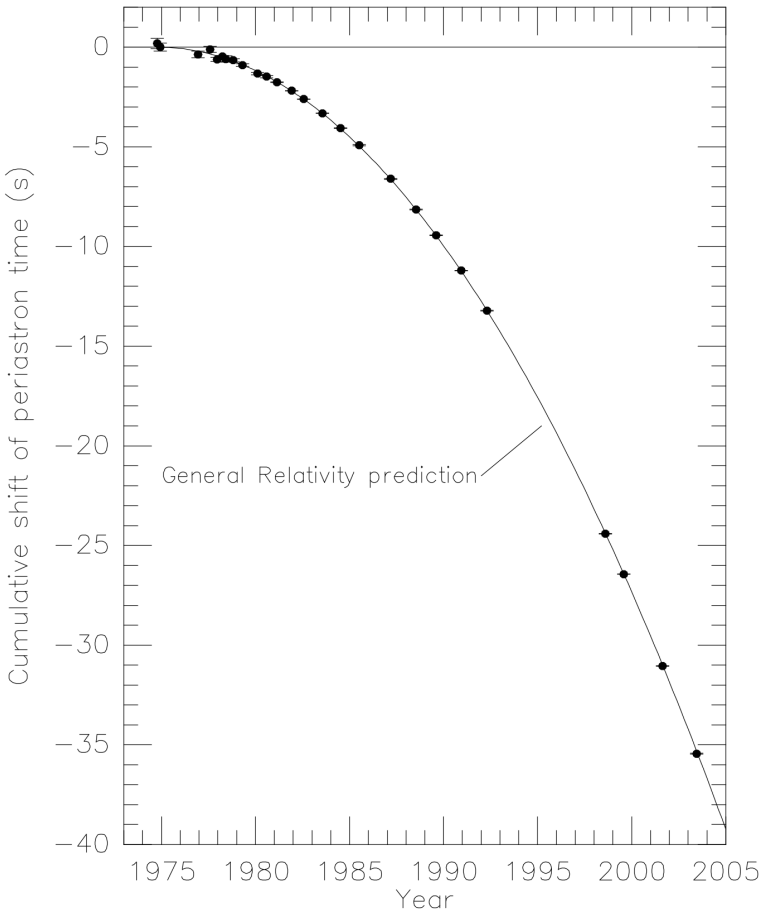

Figure 1.2: The plot of the slow shift of the periastron time from gravitational radiation reaction, aggregated over 30 years of observing the Hulse-Taylor pulsar. Plotted also is the prediction for the same system from general relativity. It is worth noting that this spectacular agreement is not a best fit line, as the parameters of the binary entirely fix the general relativity prediction. This plot is from the paper [7].

dynamics. Russel Hulse and Joseph Taylor were awarded the 1993 Nobel Prize in Physics for their observation of the Hulse-Taylor pulsar.

Construction of the Laser Interferometer Gravitational wave Observatory (LIGO) began in 1994. LIGO operates by splitting a high-intensity laser beam and sending it down a pair of 4 km-long arms, then reflecting the laser light at the ends of the arms with a pair of test mass mirrors. Once the beams return to the original separation point, they are recombined, which produces an interference pattern which measures the position of the test masses to a precision far better than one wavelength

of the light used. The original LIGO detector operated for several scientific runs (denoted S1 trough S5) intermittently from years 2002-2010. At design sensitivity, the original LIGO detectors had the ability to detect binary neutron star mergers out to 18 Mpc. While the original LIGO detectors did make any positive detections they were a crictical first step in the ultimate construction of Advanced LIGO, hereafter referred to often as simply LIGO.

The upgrades and construction of the new Advanced LIGO detector begain in 2008. By 2015, the sensitivity of the upgraded detectors had far outstripped the original LIGO detectors, and they had the ability to detect binary neutron star mergers out to $40 - 80$ Mpc [8]. An important realization for these sensitivity improvements is that the range improvements increase the detection volume by a cube of the range. A factor of two improvement in the detection range will have the practical effect of multiplying the rate of interesting detections by eight. The improved sensitivity finally led to the first direct detection of gravitational waves in September of 2015 [2], in which Advanced LIGO observed the spectacular merger of a pair of ~ 30 solar mass black holes. The obervation was an astounding accomplishment, both for the first direct confirmation of the strong-field dynamics of general relativity, and for the technological development that was necessary to obtain the measurement. The aLIGO detector extracted the data from the movement of the test masses by $\sim 10^{-19}$m, and positively confirmed the signal from a black hole merger at a distance of 420^{+150}_{-180} Mpc. Critical for the detection, though, was highly precise source modelling of the gravitational waveform made possible by a thorough understanding of the Einstein field equations and numerical relativity simulations of the possible events.

One of the important near-future developments in the budding field of gravitational wave astronomy is the construction of space based gravitational wave interferometers. The basic method of detection is similar to ground based interferometers, but the vacuum of space permits far longer detector arms by only constructing the emitter and detector corners of the array. The accepted proposal for the Laser Interferometer Space Antenna (LISA) detector positions three satellites of similar construction at the corners of an equilateral triangle, $2.5 \cdot 10^6$ km apart [9]. The symmetric design of the detector permits a detection power approximately that of a pair of detectors stacked at the same location. The comparatively noise-free environment of space and the noise reduction from careful isolation of the test masses within the satellites will ensure the detection capability to observe

numerous events inacessible to ground based detectors. Among the events made possible with space based detectors are pairs of white dwarfs, aLIGO sources early in their inspiral, supermassive black hole binaries, and solar mass objects inspiralling into supermassive black holes. Compact object binaries are classified by their mass ratio, as the ratio of the masses is most relevent for modelling the dynamics and for the characteristics of the signal. When the ratio of the larger (primary) mass to the smaller (secondary) mass m_1/m_2 is $\sim 1-10$, the binary is said to have comparable masses, when the mass ratio is $\sim 100-1000$, the source is referred to as an Intermediate Mass Ratio Inspiral (IMRI), and when the mass ratio is $\sim 10^4-10^9$, the source is referred to as an Extreme Mass Ratio Inspiral.

The sensitity of LISA should permit detection of $\sim 2-2000$ gravitational wave events from EMRI events, which are caused by solar mass objects inspiralling into supermassive black holes. These events are a particularly exciting source for space-based detectors, as the small object lingers for many orbits in the strong-field region of the supermassive black hole. The long inspiral will give rise to a long, clear signal which encodes detailed information about the gravitational field and dynamics near supermassive black holes. Such detections will give rise to important tests of general relativity, as well as provide interesting data about the population of supermassive black holes in the LISA detection volume. These ambitious science goals require, in addition to the vast engineering accomplishment of the satellites themselves, significant development in the modelling of sources which will permit the extraction of the long EMRI signals from the LISA data. The ongoing effort to model the dynamics of EMRIs is the focus of self force research, for which I give an overview in Section 1.3, and is the focus of the advancements for EMRI source calculations presented in Chapters 4, 5, and 6 of this dissertation.

1.2 Introduction to gravitational waves

1.2.1 Gravitational waves in a Minkowski background

Before moving to the more involved treatments of gravitational radiation, I review the textbook treatment of plane propagating gravitational waves in a flat spacetime [10–12]. The background

metric for this simple derivation is taken to be Minkowski,

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2. \tag{1.1}$$

We now study small perturbations about this background, such that the full metric can be written as

$$g_{\mu\nu} = \eta_{\mu\nu} + \varepsilon h_{\mu\nu}, \tag{1.2}$$

where η is the background Minkowski metric and h is the small perturbation about that background. The small metric perturbation inherits a large space of gauge freedom from full diffeomorphism freedom of the metric. The linearized, order ε gauge transformation of the metric perturbation can be written as,

$$h'_{\mu\nu} = h_{\mu\nu} + \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu, \tag{1.3}$$

where ξ is an arbitrary vector field, ∇_μ is the covariant derivative associated with the background spacetime, and h' will produce identical physical predictions to h, up to errors of order ε^2. In flat vacuum spacetime, the gauge symmetry implies that there are only two physically independent degrees of freedom encoded in the metric perturbations.

I now take advantage of the gauge freedom of the metric perturbation to impose convenient choices for the metric perturbation for computation in the Einstein field equation. I take the common choice of the Lorenz, or harmonic, gauge,

$$\nabla^\mu \left(h'_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h'^{\lambda}{}_{\lambda} \right) = 0. \tag{1.4}$$

To see that this gauge may be applied in general, consider the case in which we start with a metric perturbation $h_{\mu\nu}$ which does not yet satisfy the Lorenz gauge condition. I then wish to show that there will always exist a vector field ξ_μ, such that the gauge-transformed field (1.3) obeys the Lorenz gauge condition (1.4). Starting from an arbitrary metric perturbation $h_{\mu\nu}$, the harmonic gauge condition implies the requirement on the gauge vector,

$$\nabla^\mu \nabla_\mu \xi_\nu = \nabla^\mu \left(h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h^{\lambda}{}_{\lambda} \right). \tag{1.5}$$

This, however, is simply the wave equation for a vector field in flat spacetime, which possesses solutions for arbitrary, sufficiently well-behaved sources. Therefore, there generically exists a gauge vector ξ which will bring any metric perturbation $h_{\mu\nu}$ to a Lorenz gauge perturbation $h'_{\mu\nu}$ [10].

There is a residual freedom in the metric perturbation $h'_{\mu\nu}$, as we may yet transform by any vector field which satisfies the homogeneous wave equation $\nabla^\mu \nabla_\mu \xi_\nu = 0$ without disrupting the Lorenz gauge. We may then use the residual gauge to demand that the timelike and trace components of the metric perturbation vanish, ${h^{TT\mu}}_\mu = 0$, $h^{TT}_{t\mu} = 0$ [10].

Only the two physical degrees of freedom remain. Consider a wave which propagates along the $\hat{z}$ direction. We write the resulting plane wave mode as

$$h_{\mu\nu} = e^+_{\mu\nu} h_+ e^{ik_z z - i\omega t} + e^\times_{\mu\nu} h_\times e^{ik_z z - i\omega t}, \tag{1.6}$$

where e^+ and $e^\times$ are the 'plus' and 'cross' polarization modes and $\omega/k_z = c$. These modes can be written as matrices

$$e^+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad e^\times = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.7}$$

From this form, the inherent quadrupolar ('spin-2') nature of a gravitational wave becomes immediately apparent. Consider a rotation about the z-axis: a rotation by any multiple of π will map the solution onto itself, for either mode of the radiation. This is to be contrasted with the electromagnetic polarization modes which are invariant only under discrete rotations by multiples of 2π.

As a gravitational wave passes a set of observers, the result is a slight alteration of the measured distances along the axes orthogonal to the wave's direction of propagation. The measured distance between freely falling observers may be approximated using the geodesic deviation n^α, which evolves according to:

$$(u^\mu \nabla_\mu)^2 n^\alpha = -{R^\alpha}_{\nu\lambda\rho} u^\nu n^\lambda u^\rho, \tag{1.8}$$

where ${R^\mu}_{\nu\lambda\rho}$ is the Riemann tensor, which contains all of the curvature information described by the metric tensor, and u^λ is the velocity tangent to the worldline in question. In the particular case of a weak plane gravitational wave propagating along the $\hat{z}$ direction in a Minkowski background and a geodesic deviation initially in the xy-plane $n^\mu = \{0, n^x, n^y, 0\}$, the geodesic deviation equation

becomes

$$\partial_t^2 n^x = \frac{1}{2} h_+ n^x \omega^2 e^{ik_z z - i\omega t} + \frac{1}{2} h_\times n^y \omega^2 e^{ik_z z - i\omega t}, \tag{1.9a}$$

$$\partial_t^2 n^y = -\frac{1}{2} h_+ n^y \omega^2 e^{ik_z z - i\omega t} + \frac{1}{2} h_\times n^x \omega^2 e^{ik_z z - i\omega t}. \tag{1.9b}$$

Taking advantage of the construction that the oscillations will be very small deviations from the original lengths $n^x(0)$ and $n^y(0)$, the solutions are then:

$$n^x \approx n^x(0) - \frac{1}{2} h_+ n^x(0) e^{ik_z z - i\omega t} - \frac{1}{2} h_\times n^y(0) e^{ik_z z - i\omega t} \tag{1.10a}$$

$$n^y \approx n^y(0) + \frac{1}{2} h_+ n^y(0) e^{ik_z z - i\omega t} - \frac{1}{2} h_\times n^x(0) e^{ik_z z - i\omega t}, \tag{1.10b}$$

For freely falling observers, then, the passage of a gravitational wave may be detected by measuring the time of propagation of a reliable signal, such as laser pulses, between observers [13]. Alternatively, for bound objects, the effect of a gravitational wave passage will induce a small force (by the equivalence principle) orthogonal to the direction of gravitational wave propagation. The small, oscillatory force might then be detected by the use of an extremely high-quality mechanical oscillator [14].

1.2.2 Frequently used methods for source modelling

In this section, I will discuss the basics of each of the main methods for approximating the dynamics of gravitational inspirals. The three main techniques that have seen the most success to date are the Post-Newtonian approximation, numerical relativity, and self force methods. The parameter space for the set of binaries can be described by the pair of dimensionless parameters v/c, where v is the characteristic velocity scale of either companion, and the ratio of the mass of the large comanion to the mass of the small companion m_1/m_2. The restriction to bound binary sources ensures that large separation of the two companions relative to their Schwarzschild radii directly corresponds to small v/c. A sketch of the parameter space and the domains of validity of the relevant approximation techniques is given in Fig. 1.3.

Post-Newtonian approximation methods benefit from being comparatively direct to compute, and hold sway in the regime where the binary separation is large. Direct simulation of the long epoch in the weak-field region is computationally expensive, both due to the large simulation domain

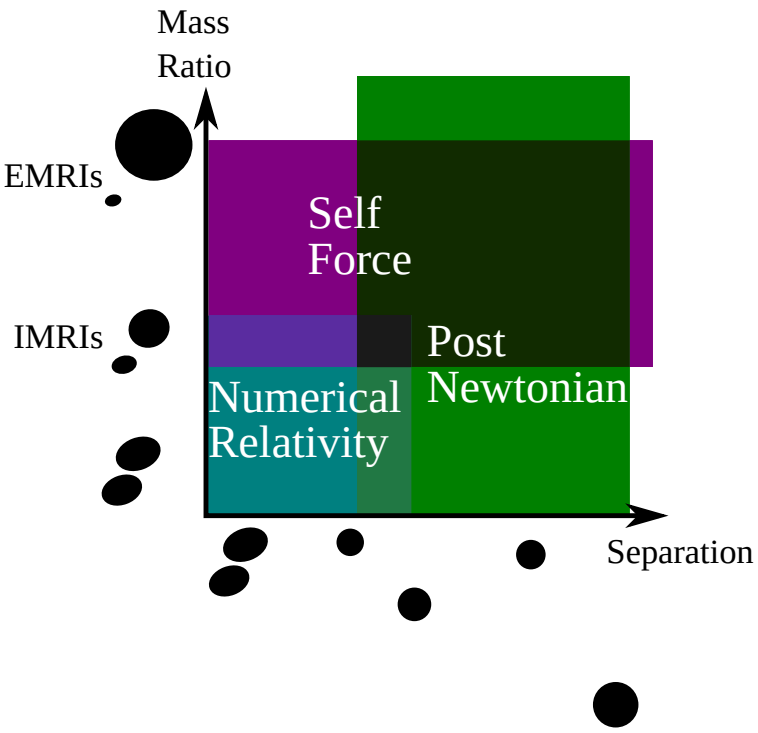

Figure 1.3: The domains of validity for the three principal methods for approximating the inspiral of binary compact objects: Post-Newtonian formalism, numerical relativity, and self force methods.

and due to the long duration of the inspiral. Instead, perturbative expansions in the sub-relativistic velocity of the binary permit the computation high-fidelity waveforms at high separations. The methods of Post-Newtonian expansions are discussed in Sec. 1.2.2.2

Numerical relativity is the currently the only method for predicting the gravitational effects where no separation of scale is available. The most important region of the parameter space which numerical relativity is required is when the separation between the bodies is similar to the characteristic curvature of spacetime, and the masses are similar. The first successful simulation of a binary black hole merger was accompished by Frans Pretorius [15], and many modern methods have been developed to advance the generality and granularity of direct simulations. Numerical relativity methods are discussed in somewhat more detail in Sec. 1.2.2.2.

Finally, the most relevant technique to the discussions in this dissertation is the self force method. This technique is the most useful in the regime where the mass ratio of the binary is large, yet the small companion is permitted to move into the strong field region and acquire a relativistic velocity. This regime also permits a perturbative expansion in the small parameter m_2/m_1, but

involves computations of the interactions in the strong-field region of the large companion. Self force techniques are introduced in Sec. 1.2.2.3, and discussed in more detail in Sec. 1.3

The combination of Numerical relativity and Post-Newtonian waveform predictions have proved critical for the modern detection of gravitational waves. As I will discuss in Chapter 2, the processing of candidate gravitational wave detections relies on precise predictions of the waveform. The long lead-up to merger can be computed by Post-Newtonian methods, but the final and loudest several orbits require numerical relativity. Due to the large separation of scales, direct simulation cannot handle high mass ratio binaries. The corresponding precise waveform predictions for high mass ratio binaries will require self force techniques which can capture the phase of the waveform to sub-radian precision. In Fig. 1.3, I have sketched the parameter space relevant to each of the three main techniques.

1.2.2.1 Post-Newtonian approximations

Much of this brief summary of Post-Newtonian approximation methods is based on the excellent review article by Luc Blanchet [16] and references therein. The Post-Newtonian approximation is computed via an iterative technique. First, well-understood Newtonian potentials are used to compute the motion of the binaries. Next, the motion is used to derive small corrections to the metric tensor beyond the leading Newtonian approximation. These small corrections, in turn can be fed back into the motion, and so on.

Formally, the Post-Newtonian approximation expands the metric perturbation in powers of $1/c$, with appropriate relative scaling of the time and space components of the metric [16, 17]:

$$g_{tt} = -c^2(1 + c^{-2}h^{(0)}_{tt} + c^{-4}h^{(1)}_{tt} + \mathcal{O}(c^{-5})) \tag{1.11a}$$

$$g_{ti} = c(c^{-3}h^{(1)}_{ti} + \mathcal{O}(c^{-4})) \tag{1.11b}$$

$$g_{ij} = (\delta_{ij} + c^{-2}h^{(1)}_{ij} + \mathcal{O}(c^{-3})), \tag{1.11c}$$

where in Post-Newtonian expansions, the perturbation quantities are typically given alternative notation, $h^{(0)}_{tt} = 2\Phi$, $h^{(1)}_{tt} = 2(\Phi^2 + \psi)$, and $h^{(1)}_{ti} = 2\zeta_i$. It is notable that in performing an expansion in terms of dimensionful quantities, we are implicitly assuming that all factors of velocity that might appear in the metric perturbation are individually small compared to the speed of light. This

assumption is well-justified for compact sources at high separation, but is an important limit on the applicability of Post-Newtonian expansions.

The effective stress energy tensor, which includes nonlinear contributions from the metric perturbation, is similarly expanded

$$T_{(\text{eff})\alpha\beta} = \sum_{m=-4}^{\infty} \frac{1}{c^m} T^{(m)}_{(\text{eff})\alpha\beta}(\vec{x}, t). \tag{1.12}$$

From the expansions (1.11) and (1.12), each subsequent order of the metric perturbations may be derived from the Poisson equations,

$$\nabla^2 h^{(m)}_{(\text{eff})\alpha\beta} = 16\pi G T^{(m-4)}_{(\text{eff})\alpha\beta} + \partial_t^2 h^{(m-2)}_{\alpha\beta}. \tag{1.13}$$

At subleading orders in this expansion, the field dependent source is of non-compact support; it has nonvanishing value at all values of r. As a result of the instantaneously-transmitting Poisson equation, the extended source reaches to $r \to \infty$. The infinitely large extended source then gives rise to a divergent solution to the Poisson equation at subleading order. In order to preserve the approximation method, the Poisson equation must be solved on a domain with outer radius $R \ll cT$, where T is the characteristic timescale of the evolution. The solution to the problem at large scales is to only use the Post-Newtonian approximation in a finite subregion near the binary, and match to a Post-Minkowski approximation at distances $r \sim cT$.

The equations of motion of the matter sources can be computed using the geodesic equation evaluated on an appropriately regularized form of the perturbed metric. The resulting acceleration follows the Newtonian gravitational force law at leading order, and acquires first corrections at post-1-Newtonian order, which is suppressed by a factor of v^2/c^2. Dissipative acceleration effects first arise at Post-2.5-newtonian order, suppressed compared to the leading Newtonian effects by v^5/c^5.

There is an important overlap in the parameter spaces of post Newtonian approximations and gravitational self force computations. When both the separation and the mass ratio is large, both approximations are valid, and in this region the two approximation methods can be used as checks to one another. The comparison of the equations of motion has been carried out to leading order in the mass ratio, and through 7.5 Post-Newtonian order [18, 19], with excellent agreement.

1.2.2.2 Numerical relativity

In computations of compact binary inspirals, numerical relativity is typically limited to computing the dynamics and effects of comparable mass objects in the strong field region. Of course, numerical methods are used in all approximation schemes, but here we make use of the conventional term 'numerical relativity' to describe the direct simulation of general relativity discussed below. These techniques make no analytical approxmations, and instead simulate the Einstein field equations in their full, nonlinear glory.

To accomplish this, the Einstein field equations must be cast into a form which is amenable to numerical evaluation. In particular, it is advantageous to identify a foliation of spacelike 3-dimensional hypersurfaces parameterized by some time coordinate t, which need not have any direct relation to convenient or often-used coordinates of analytic spacetimes. The normal vector n^α is then used to relate the metric to an induced 3-metric on each hypersurface

$$\gamma_{\alpha\beta} = g_{\alpha\beta} + n_\alpha n_\beta. \tag{1.14}$$

The congruence of a second, related vector field, $t^\mu = \alpha n^\mu + \beta^\mu$ joins points of equal hypersurface coordinates $\vec{x}$ and unequal time t. The parameters α and β^μ are known as the Lapse and the Shift, and are functions only of the choice of foliation and coordinates within each foliation.

The spacetime is then typically evolved [20] using the 3+1 ADM equations. Following the above discussion, the metric is written as

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ij}(dx^i + \beta^i)(dx^j + \beta^j dt) \tag{1.15}$$

The resulting formalism decomposes the Einstein field equations into purely spatial constraints which restrict initial condition choices and equations involving time derivatives which can be used to compute the metric γ_{ij} and extrinsic curvature $K_{ij} \equiv -\gamma_i{}^k \gamma_j{}^l \nabla_k n_l$ on each successive hypersurface.

Two principal methods have been developed for simulating the 3+1 equations of general relativity: finite difference methods and spectral methods. Finite difference methods introduce a numerical gridding of spacetime, and evaluate the evolution of the ADM equations by computing discrete derivatives using differences of quantities between the spacetime gridpoints. The most

highly used current code base which implements finite differencing algorithms for black hole inspirals is the Einstein Toolkit [21].

Spectral methods, by contrast, introduce a decomposition of the fields as a linear combination of mode functions. Conceptually similar to a Fourier series, spectral expansions have the desirable property that, for well-behaved functions, a spectral series converges exponentially. The current state of the art code, SpEC, takes advantage of multiple coordinate frames and expands the relevant fields in angular harmonics and Chebyshev polynomials of radial coordinates [22–24].

An exciting new method which has been introduced more recently for numerical relativity simulations is known as the Discontinuous Galerkin method. This can be conceptualized as a compromise method between spectral and finite difference methods, and offers some attractive improvements over both. The computational domain is segmented into several subdomains, and on each subdomain, the equations are decomposed into spectral modes. The method shows promising initial results in the ability to handle matter sources to high precision in an efficient and scalable way [25].

1.2.2.3 Self force in black hole perturbation theory

Self force methods are constructed to address the region of parameter space inaccessible both to Post-Newtonian approximation and numerical relativity: the case in which the mass ratio is large and the separation is similar to the curvature scale. Post-Newtonian methods are invalidated as the characteristic timescale approaches GM/c, where M is the mass of the large companion. Numerical relativity struggles to address this case, as the characteristic scale of the small companion informs the refinement necessary in the computational scheme. For finite-differencing methods, this corresponds to a prohibitively fine mesh refinement, and for spectral methods, it corresponds to a dependence on high-order modes. To make matters worse, the inspiral lasts a number of orbits which scales with the mass ratio, ensuring that the numerical relativity code would need to evolve the system over a prohibitively long radiation-reaction time.

Self force methods instead leverage the large separation of masses to define a small parameter of the system $\varepsilon = m/M$, equal to the ratio of the mass of the small companion to the mass of the large companion. For extreme mass ratio inspirals, $\varepsilon \sim 10^{-4} - 10^{-6}$. Then, at all points away from the small companion, the metric can be written as a perturbative expansion about the large

companion's stationary black hole metric.

$$g_{\alpha\beta} = g^{(0)}_{\mu\nu} + \sum_n \varepsilon^n h^{(n)}_{\mu\nu} \tag{1.16}$$

These metric perturbations $h^{(n)}$ can be treated using black hole perturbation theory, in which the various modes are derived from the wave equation in a black hole background. The machinery for computing such modes is well-established, and a number of techniques are available for Kerr black holes. The full set of components of the metric perturbation can be evolved in the Lorenz gauge, which yields a set of 10 coupled wave equations. Alternatively, for some computations, it is sufficient to derive only the Weyl scalars, which can be derived from the Teukolsky master equations.

The motion of the small companion is then computed using self force theory. The derivation of this motion is a subtle process, as the metric perturbations which govern the motion diverge at the position of the small companion. The corrections to the motion are derived via regularization or matched asymptotic expansions. The main results from the first and second order self force derivations are that the motion of the small object is geodesic in an appropriately defined, regular perturbed metric.

Due to the central importance of the self force techniques, I will discuss the reasoning in detail and the most salient prior results in a dedicated section below (Sec. 1.3).

1.2.3 An illustration of basic source approximations

In this section, I discuss an illustrative example of an approximation for the generation of gravitational radiation. The approximation I use follows the frequently used strategy of first computing the approximate motion of a compact binary inspiral, then using the full approximate worldline to compute the radiation emitted by the source. This general strategy is employed in far more elaborate approximations, so the brief discussion here is a reasonable prototype for the more thorough computations discussed in later chapters. The basic formalisms I use in this section are the Post-Newtonian formalism [10–12], and the quadrupole approximation of a gravitationally radiating system [26].

The field equations of general relativity ensure the conservation of stress-energy $\nabla_\mu T^{\mu\nu} = 0$, which directly implies the relativistic kinematic laws of conservation of 4-momentum, as well as

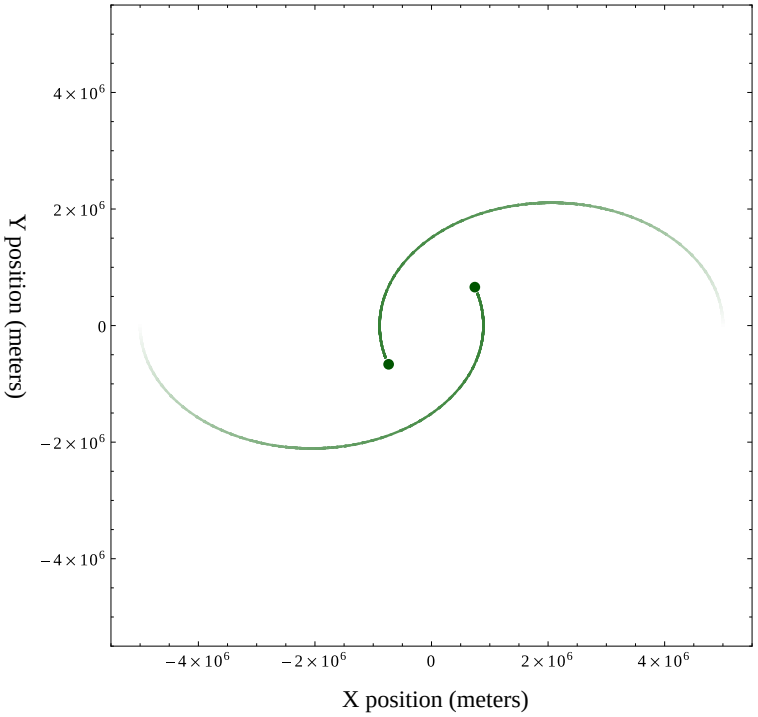

Figure 1.4: Symmetric, Post-Newtonian orbit, evolved for $t = 2$sec with a pair of $1.4M_\odot$ compact objects, and starting velocity $v = 6.6 \cdot 10^{-3}c$. The orbit has maximum velocity $v \approx .03c$, so the simple Post-Newtonian approximation should give within percent accuracy.

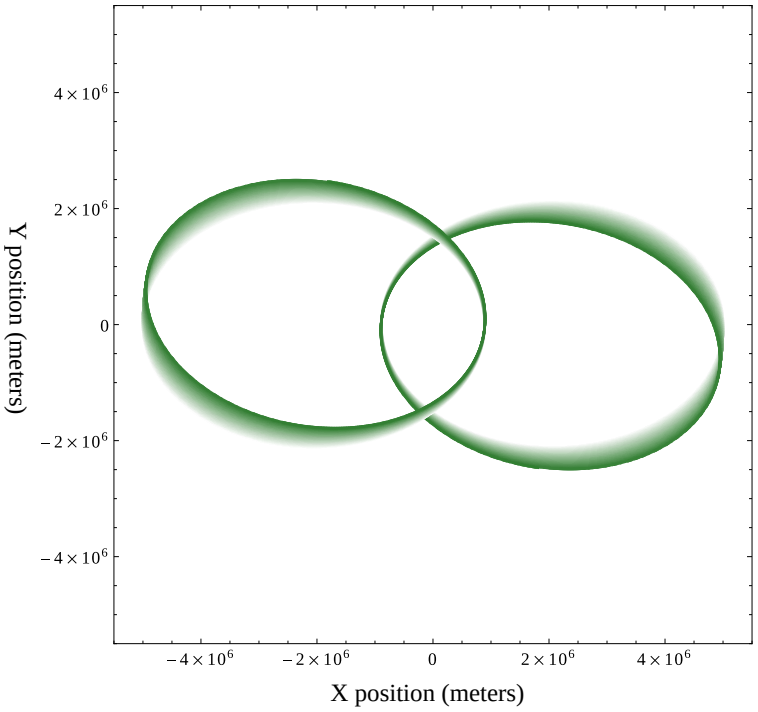

Figure 1.5: Symmetric, Post-Newtonian orbit, evolved for about 25 orbits with a pair of $1.4M_\odot$ compact objects, and starting velocity $v = 6.6 \cdot 10^(- 3)c$, to show the relativistic correction in the precession of periapse.

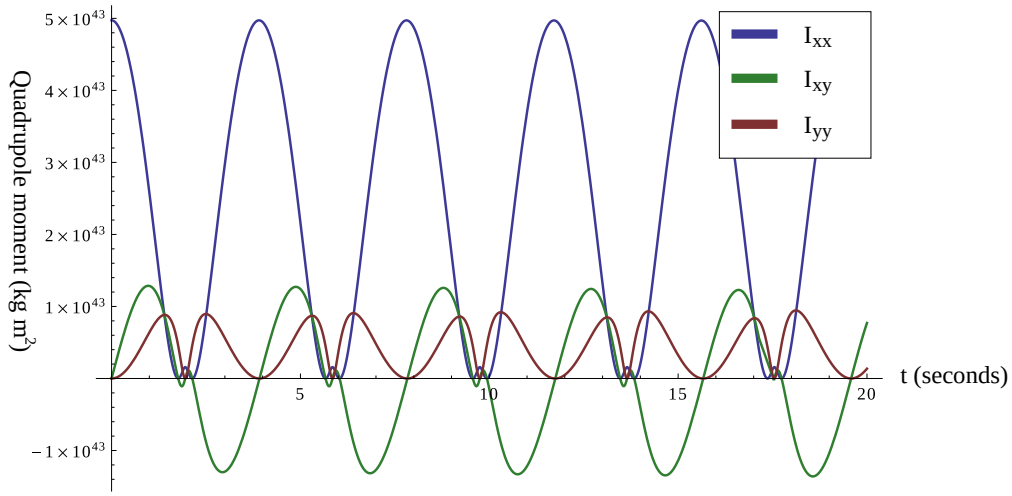

Figure 1.6: The resulting quadrupole moment of the PN binary as a function of time. As the orbit precesses, the I_{xx} component will evolve into the I_{yy} component.

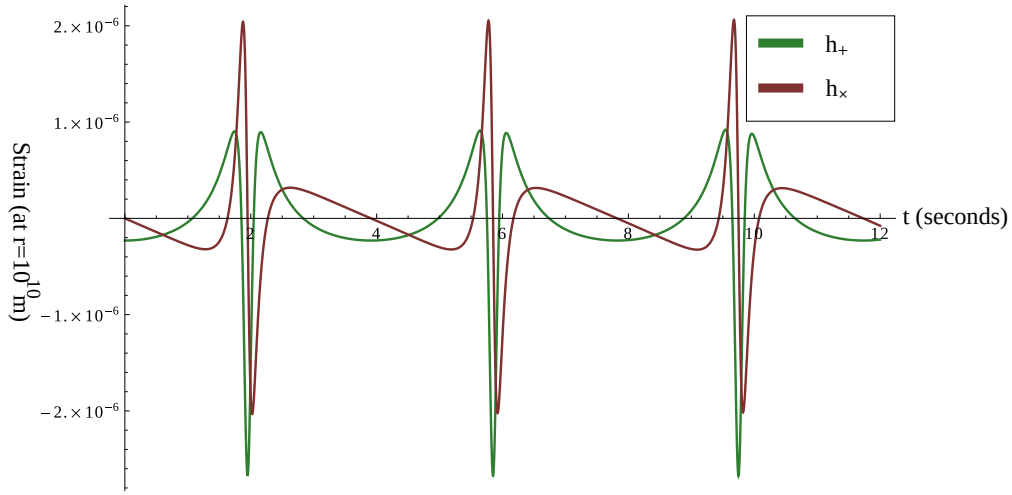

Figure 1.7: The dimensionless strain associated with the gravitational radiation emitted by the Post-Newtonian binary.

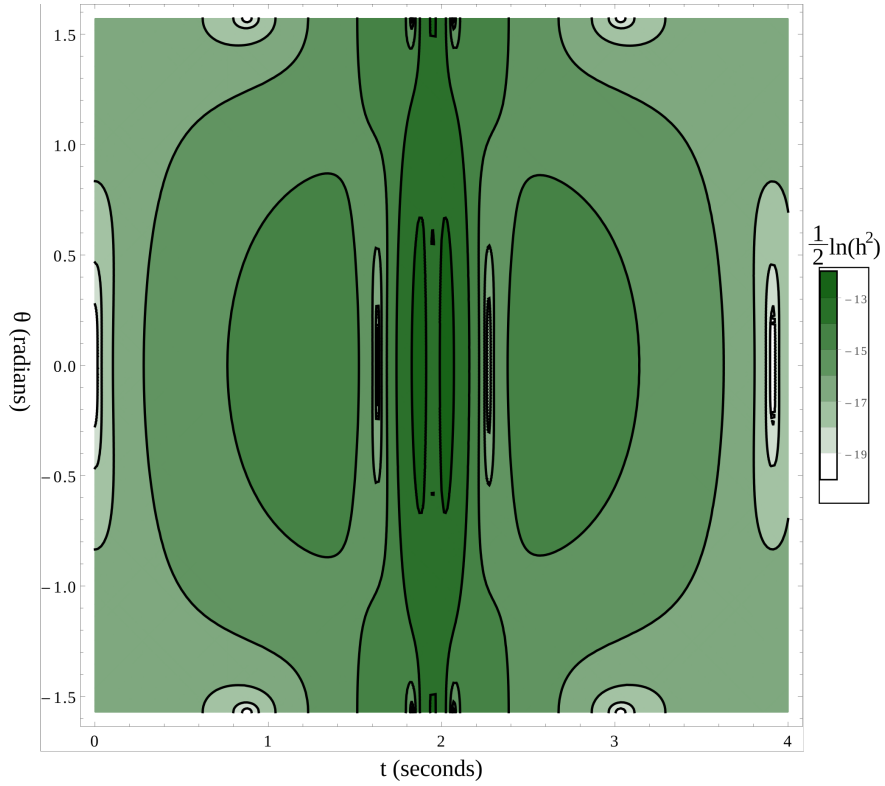

Figure 1.8: The antenna pattern of a measure of the emitted power, $(1/2)\ln(h^{\mathrm{TT}}_{\mu\nu}h^{\mathrm{TT}\mu\nu})$, as a function of time and the polar angle of the binary at $\phi = 0$. The plot is symmetric about the point of periapse due to the construction of a pair of equal masses in the center-of-mass frame.

angular momentum. The full set of conservation laws in general relativity directly implies that gravitational radiation caries no dipole moment, and the overall mass dipole of any system is conserved $\dot{P}^i = 0$. Therefore, unlike electromagnetic radiation, for which the leading mode arises from charge dipole oscillations, the leading gravitational radiation mode arises from mass quadrupole oscillations.

Due to the practical consideration that all detectable sources of gravitational waves are astrophysical in nature, I consider for this illustrative argument a binary system of compact objects. Here, and throughout this dissertation, I use the astrophysical term 'compact' to refer to objects for which their characteristic size is similar in magnitude to the characteristic scale of the spacetime curvature which they source. Practically speaking, this definition of compact includes black holes and neutron stars, but excludes most other material bodies, such as main-sequence stars and planets.

For the simplicity of this basic treatment, I assume an equal-mass binary in a elliptical orbit, with separation r_0 far greater than the objects' characteristic size R, $r_0 \gg R$. Due to the large separation, the binary has sub-relativistic velocities $v \ll c$, so I make use of the first couple of terms from the Post-Newtonian framework to derive the orbit. A more detailed discussion of modern advancements in the application of the Post-Newtonian approximation methods to precise derivation of gravitational waveforms is given in Subsection 1.2.2.1. To get a basic basic idea of the formalism, it is sufficient to note the initial expansion of the metric tensor in the Post-Newtonian framework.

$$g_{00} = -2\phi^2 - 2\psi, \tag{1.17a}$$

$$g_{i0} = \zeta_i, \tag{1.17b}$$

$$g_{ij} = -2\delta_{ij}\phi. \tag{1.17c}$$

The Post-Newtonian technique proceeds by finding the leading-order (Newtonian) field ϕ from the sources. The leading field is then used to determine the leading order motion of the sources, which in turn is used to derive the corrections ζ and ψ to the Post-Newtonian metric. The Post-Newtonian field contributions can then be used to derive corrections to the motion, and so on.

For my illustrative example, I'll carry out the derivation of the motion only to first Post-Newtonian order. Each order of the approximation is suppressed by a factor of v^2/c^2. It is a well-known result that proceeding directly with the Post-Newtonian expansion will find gravitational radiation effects only at post-5/2 Newtonian order. Instead of carrying the approximation to this high order, I will simply introduce the quadrupole approximation for the radiation from a gravitating system to illustrate a coarse-grained version of the resulting wave dynamics.

For brevity, I simply quote the results of a Post-Newtonian expansion of a symmetric two-body quasicircular source, obtained from the high-order Post-Newtonian expansions aggregated by [16]. The acceleration is (imposing $m = m_1 = m_2$ and $x = x_1 = -x_2$),

$$\vec{a}_1 = -\frac{Gm}{2r^2}\hat{r} + \frac{1}{c^2}\left[\frac{9G^2m^2}{8r^3}\hat{r} + \frac{3Gm}{8r^2}(\vec{v}\cdot\hat{r})\hat{r} + \frac{Gm}{4r^2}(\vec{v}\cdot\hat{r})\vec{v}\right]. \tag{1.18}$$

This is a simple second-order differential equation that can be used directly to find the orbit as a function of time. For the illustrative example, I have plotted the resulting orbits from a pair of

solar-mass compact objects, with starting parameters, $m = 1.4M_\odot$, $r_0 = 5 \cdot 10^6$m, $v_0^\phi = 6.67 \cdot 10^{-3}c$, in Fig. 1.4, 1.5.

At leading order, and in the Lorenz gauge, the wave equation for the trace-reversed metric perturbation $\bar{h}$ is

$$\Box \bar{h}^{\mu\nu} = -16\pi T^{\mu\nu}. \tag{1.19}$$

Using, then, the flat space Green's function in the slow-motion approximation, the metric perturbation far from the source can be approximated as

$$\bar{h}^{\mu\nu}(t,x) = \frac{4}{r}\int T^{\mu\nu}(x', t-r')d^3x' \tag{1.20}$$

Noting that in the Post-Newtonian slow-motion approximation, the T^{00} component is dominant, and expanding the argument of the integral near $x' = 0$, we find that the dominant component of the metric perturbations is

$$\bar{h}^{jk}(t,x) = \frac{2}{r}\int d^3x' \frac{\partial^2}{\partial t^2} T^{00}(t, t-r')x'^j x'^k \equiv \frac{2}{r}\frac{d^2 I^{jk}(t-r)}{dt^2}, \tag{1.21}$$

where I_{jk} is the quadrupole moment of the source. Transforming this to the transverse-traceless components described in Section 1.2.1, we obtain

$$h_{jk}^{TT}(t,x) = \frac{2}{r}\frac{d^2}{dt^2}(P_{jl}P_{km}I^{lm} - \frac{1}{2}P_{jk}P_{lm}I^{lm}) \tag{1.22}$$

I therefore apply the quadrupole formula (1.21) to the low-order post Newtonian source obtained from evolving the post-1-Newtonian equation of motion (1.18), shown in Fig. 1.5. The quadrupole moment for the symmetric source is easy to derive by treating the pair of masses as point sources, which gives

$$I_{xx} = mr(t)^2\cos^2(\phi(t)) \qquad I_{xy} = mr(t)^2\cos(\phi(t))\sin(\phi(t)) \tag{1.23a}$$

$$I_{yx} = mr(t)^2\cos(\phi(t))\sin(\phi(t)) \qquad I_{yy} = mr(t)^2\sin^2(\phi(t)). \tag{1.23b}$$

I have plotted the resulting time-dependent quadrupole moment, and the gravitational wave signal along the z-axis of the spacetime in Figs 1.6, 1.7. Finally, the quadrupole nature of the waves ensures that there is a nontrivial "antenna pattern" for the strength of emission in the two polarizations. I have plotted the relative strengths of gravitational wave emission over the sphere in Fig. 1.8.

I emphasize that this is the simplest, coarsest description of gravitational waves which still captures some general relativistic effects. The substantial industry of evaluating binary inspirals with great precision uses a combination of sophisticated techniques which I describe broadly in Subsection 1.2.2.

1.3 | self force methods for accurate EMRI predictions

Space-based gravitational wave detectors, such as LISA, are expected to have sufficient sensitivity to detect the gravitational waves emitted from 2-2000 sources with mass ratio m/M in the range $10^{-4}-10^{-9}$ per year (extreme mass ratio inspirals) . Importantly, the capability of reliable detection and parameter estimation for such systems requires precise predictions of the extreme mass ratio inspiral (EMRI) gravitational wave signals [27] (discussed in more detail in Section 2.3 below). Such systems are prohibitive for direct simulation (Section 1.2.2.2) or Post-Newtonian approaches (Section 1.2.2.1). Instead, a perturbative framework which leverages the large mass ratio is required for a precise waveform model, which requires detailed modelling of the radiation reaction effects.

The goal of the self force community is a method for efficiently computing extreme mass ratio inspirals such that the error in the phase of the waveform is small throughout the inspiral. The sub-radian accuracy goal requires a careful treatment of the metric near the massive black hole primary companion, as well as detailed backreaction of the perturbations sourced by the small secondary companion, which produce fine frequency shifts and subtle dissipative effects. These must all be accounted for to achieve the sub-radian accuracy goal and to generate trustworthy waveforms to prepare for the task of LISA data analysis discussed in section 2.3.3.

In this introduction section, I summarize the current landscape of self force derivation and computational methods. The past work of the self force community is an important foundation for the work presented in this dissertation, as much of my work builds directly on past self force derivations (presented in Chapter 3), and extends past methods for assembling a robust technique of handling the gravitational perturbations at all requisite scales to compute highly accurate EMRI waveforms (presented in Chapters 4,5, and 6).

In subsection 1.3.1, I summarize the basic mathematical structure used in self force computations. I then describe in subsection 1.3.2 the original derivation of radiation reaction effects at first

order in the charge (for electromagnetic effects), or mass (for gravitational effects). I then summarize in subsection 1.3.3 the recent derivations of the analytic form of the second order self force for gravitational effects. There are a number of procedures developed by the self force community for obtaining the resulting self forced worldline and waveform from the inspirals. I will the worldline techniques in subsection 1.3.4 and the techniques for waveform generation in 1.3.5.

1.3.1 | Overview of perturbative framework

The central construction for self force computations is the use of the mass ratio of EMRIs as a small parameter for a perturbative calculation. In particular, we use $\varepsilon \equiv m/M$, with m the mass of the small companion, and M the mass of the large companion, as a single parameter for both the expansion of the spacetime metric and the inspiralling worldline.

The metric in the vicinity of the massive compact object is assumed to have the form

$$g_{\mu\nu}(x^\mu, \varepsilon) = g^{(0)}_{\mu\nu}(x^\mu) + \varepsilon h^{(1)}_{\mu\nu}(x^\mu) + \varepsilon^2 h^{(2)}_{\mu\nu}(x^\mu) + \mathcal{O}(\varepsilon^3), \tag{1.24}$$

where $g^{(0)}_{\mu\nu}$ is the ε-independent metric associated with the larger central object, typically taken to be either a Schwarzschild or Kerr metric. The motion of the small companion in this background metric is described an ε-dependent worldline, $z^\mu(\tau, \varepsilon)$, which tracks the center-of mass motion of the small object [1]. Self force computations either expand the worldline perturbatively about a fiducial $\varepsilon = 0$ geodesic worldline [28, 29], or use the full worldline as a direct functional dependence of the metric perturbation [30].

The principal benefit to the perturbative expansion of the spacetime metric is the resulting tractability of the Einstein field equation for the metric perturbation at each successive order. Under the perturbation theory assumption (1.24), the field equation for each successive order is a linear differential equation, with a source which depends on a nonlinear combination of field quantities

[1] The center of mass of an object in general relativity is a delicate subject, and depends on various definitions and choices for the moments of the stress energy and spin supplementary condition. See the discussion of body parameters in (chapter 3) for a more complete presentation.

determined at lower orders,

$$G_{\mu\nu}[h^{(1)}] = 8\pi T_{\mu\nu}, \tag{1.25a}$$

$$G_{\mu\nu}[h^{(2)}] = -G^{(1)}_{\mu\nu}[h^{(1)}, h^{(1)}] + 8\pi T^{(1)}_{\mu\nu}[h^{(1)}], \tag{1.25b}$$

$$G_{\mu\nu}[h^{(3)}] = -G^{(2)}_{\mu\nu}[h^{(1)}, h^{(1)}, h^{(1)}] - G^{(1)}_{\mu\nu}[h^{(2)}, h^{(1)}] + 8\pi T^{(1)}_{\mu\nu}[h^{(2)}] + 8\pi T^{(2)}_{\mu\nu}[h^{(1)}, h^{(1)}]. \tag{1.25c}$$

In these expressions, the subleading contributions $G^{(n)}$ and $T^{(n)}$ denote the nth power of dependence on the subleading metric perturbation. For stress energy tensors which do not directly depend on the metric, $T^{(n)}|_{n>0} = 0$. The linearized Einstein field operator may then be further simplified by using the Lorenz gauge $\nabla_\mu(h^{\mu\nu} - (1/2)g^{\mu\nu}h^\lambda{}_\lambda) = 0$ to obtain a more tractable wave operator, or by making use of the separable scalar wave equations: Regge-Wheeler/Zerilli for Schwarzschild or Teukolsky for Kerr.

The radiation-reaction force is then determined from the metric perturbations and is used to modify to the geodesic equation of motion to evolve the worldline, backreacting to the surrounding spacetime via the perturbative Einstein field equations (1.25). This simple description captures the broad picture of the computational program, however, as I describe in the subsequent sections, the detailed computation requires a great deal of precise reasoning to obtain a reliable waveform. The singular nature of the metric near the small companion's worldline and the stringent requirement of sub-radian accuracy ensures that the computation of EMRI waveforms remains a compelling and challenging problem in modern general relativity.

1.3.2 First order self force derivations

1.3.2.1 Electromagnetism: Abraham, Lorentz, Dirac self force

The earliest derivation of the radiation-reaction force on a charged particle, accelerated by an external electromagnetic field, was performed by Lorentz [31], followed by Abraham and Dirac [32,33]. The basic problem can be explained via the simple realization from classical field theory that accelerating charges will emit electromagnetic radiation, which carries energy and momentum. The emission from accelerated charges is a phenomenon often observed in the x-rays emitted by charged particles in synchrotron storage rings as they are accelerated by strong magnetic fields. Energy and momentum conservation then demand that the charge must experience a force associated with the

radiation, decreasing its own energy and momentum by a commensurate amount. The question of the true motion of the charged body under sustained radiation-reaction effects is subtle. One should expect that the charge will experience a force both due to the external field and due to the radiation. The radiation-reaction force, or self force, is the direct result of the accelerating charge interacting with its own field.

In this section, I review the original reasoning for first order electromagnetic self force. For a discussion of a modern derivation [34], and a complete presentation of the first derivation of the second order electromagnetic self force performed by me in collaboration with my advisor, Éanna Flanagan, see Chapter 3.

The textbook derivation [35, 36] of the radiation-reaction force invokes conservation of energy and the power radiated by an accelerating particle, known as the Larmor radiation:

$$P = \frac{2}{3}\frac{e^2}{c^3}\dot{v}^2, \tag{1.26}$$

where e is the charge, c is the speed of light, v is the particle velocity, and the formula is valid to leading order in the ratio of characteristic length and timescales of the charge to the characteristic scale of the externally forced acceleration.

The net energy lost over an interval from t_1 to t_2 of a full period of a periodic motion is then,

$$\int_{t_1}^{t_2} \vec{F}_{\rm rad} \cdot \vec{v} dt = -\int_{t_1}^{t_2} \frac{2}{3}\frac{e^2}{c^3}|\dot{v}|^2 dt, \tag{1.27}$$

where $F_{\rm rad}$ is the force due to radiation emission. Performing an integration by parts, and moving the resulting acceleration term to the left-hand side of the equation, one obtains the simple force result,

$$\int_{t_1}^{t_2} \left(\vec{F}_{\rm rad} - \frac{2}{3}\frac{e^2}{c^3}\ddot{\vec{v}}\right) \cdot \vec{v} dt = 0, \tag{1.28}$$

which gives the famous Abaraham-Lorentz-Dirac self force,

$$\vec{F}_{\rm rad} = \frac{2}{3}\frac{e^2}{c^3}\dot{\vec{a}}. \tag{1.29}$$

There are a number of objections that have been raised to this expression over the years - it is an expression that yields a third-order differential equation for the position of the particle if the radiation force is then used to determine acceleration. Naive application of the formula under that

interpretation leads to pathological behavior such as runaway solutions or pre-acceleration. Any such result, however, rapidly exits the domain of validity of the approximation, invalidating the solution as a reasonable description of the dynamics.

The reduction-of-order procedure may be used to re-write the formula in terms of the external force, $\vec{F}_{\text{ext}}$ which gave rise to the motion in the first place. The reduced order equation of motion can then be written

$$m\vec{a} = \vec{F}_{\text{ext}} + \frac{2}{3}\frac{me^2}{c^3}\left[\frac{\partial \vec{F}_{\text{ext}}}{\partial t} + (\vec{v}\cdot\vec{\nabla})\vec{F}_{\text{ext}}\right], \tag{1.30}$$

which possesses no unusual pathologies. I discuss further the reduction of order procedure and its support from perturbation theory during the presentation in chapter 3.

1.3.2.2 Gravitational self force: MiSaTaQuWa formula

A similar radiation-reaction force arises in general relativity. The argument for general relativity is more subtle, though nonetheless, a similar argument may be made: gravitational perturbations may be related to local changes in the metric, and freefalling massive objects receive nontrivial corrections to their motion due to the emission of gravitational waves. Asymptotically, those waves may be found to carry energy and momentum, and a commensurate quantity of energy and momentum is lost by the emitting system.

The first order gravitational self force expression, in terms of the metric perturbation, is referred to as the MiSaTaQuWa self force, after the pair of derivations by Mino, Sasaki, and Tanaka [37] and Quinn and Wald [38]. Due to the comparative complexity of gravitational self force derivations, I will give only a brief summary in this section.

The derivation by Mino, Sasaki, and Tanaka [37] presented two methods of obtaining the first order gravitational self force. The first method bears close relation to methods used for matter-based self forces. In particular, the authors define a conserved effective stress energy tensor,

$$\mathcal{T}^{\mu\nu}[h^{(1)}, z] = T^{\mu\nu}[z] + \varepsilon T^{(1)\mu\nu}[h^{(1)}, z] - \frac{\varepsilon}{8\pi}G^{(2)\mu\nu}[h^{(1)}, h^{(1)}]. \tag{1.31}$$

The authors apply stress-energy conservation, along with the Hadamard form of the Lorenz gauge Green's function to derive the self force in terms of the the regular part of the metric perturbation.

Mino et. al. [37] also presented a method based on matched asymptotic expansions, which has since become a frequently applied technique in self force arguments. The method of matched asymptotics assumes a complimentary pair of domains, and performs an expansion of the metric perturbation tuned separately to each of the domains. In the case of a small body in a long-curvature background, the relevant expansion far from the small body ($r \gg m$) is [27]

$$g_{\mu\nu} = g^{(0)}_{\mu\nu}(x^\mu) + \varepsilon h^{(1)}_{\mu\nu}(x^\mu) + \mathcal{O}(\varepsilon^\in). \tag{1.32}$$

Meanwhile, very near the small companion ($r \ll M$), the metric possesses a well-organized expansion in the scaled coordinates $\{x^i/\varepsilon, t\}$,

$$g_{\mu\nu} = g^{(\text{body})}_{\mu\nu}(x^i/\varepsilon, t) + \varepsilon H^{(1)}_{\mu\nu}(x^\mu/\varepsilon) + \mathcal{O}(\varepsilon^2). \tag{1.33}$$

where the metric $g^{(\text{body})}_{\mu\nu}$ is constructed by demanding that the metric of the small object sufficiently resemble the Schwarzschild metric. A more careful construction yields a covariant 'puncture' metric, first derived in [39].

The matched asymptotic expansions may then be expanded in terms of distance from the small companion's worldline in the common region of validity. The interior metric is expanded in powers of the inverse radius ε/r, while the exterior metric is expanded in powers of r. Terms are then matched order-by-order to determine the homogeneous modes of the interior spacetime, and the results of the small companion's presence on the external spacetime. The consistency with the matched asymptotic expansion determines then the coordinate (gauge) transformation between the interior spacetime, in which the small companion is fixed and centered, and the external spacetime, in which the large companion is fixed and centered.

The final result of both the derivation by Mino, Sasaki, and Tanaka and that of Quinn and Wald is the first order self force,

$$\frac{Du^\mu}{D\tau} = (g^{\mu\nu} + u^\mu u^\nu)(2h^R_{\nu\lambda;\rho} - h^R_{\lambda\rho;\nu})u^\lambda u^\rho, \tag{1.34}$$

where h^R is the regular part of the self field obtained from integrating only over the extended 'tail' of the Green's function.

The derivation by Quinn and Wald [38] follows a more formal (axiomatic) version the conserved stress-energy method by Mino, Sasaki, and Tanaka, but arrives to the additional important real-

ization that, in order to compute just the dissipative part of the self force, the regularized field which appears in the self force formula can be represented as the half-retarded plus half-advanced self field, also referred to as the radiative self field.

While the radiative self field is explicitly regular, it the has the undesirable property of acausal dependence [27]. The direct dependence on the advanced self field ensures that the 'tail' contribution depends on the state of the worldline arbitrarily far in the future, even for finite displacements from the worldline. Detweiler and Whiting [40] gave a reformulation of the regular field $h^R_{\mu\nu}$ which depends only on source points that are either spacelike separated from or in the timelike past of the field point. In addition, the Detweiler-Whiting regular field may be used directly in the self force formula (1.34), and recoovers the full self force, including conservative effects.

1.3.3 Second order self force derivation

In this section, I restrict attention to the two past derivations of second order gravitational self force by Adam Pound [41, 42] and by Samuel Gralla [43]. In chapter 3, I give the details of my own first derivation for the second order self force for matter fields.

Both currently published second order gravitational self force derivations [41–43] apply the techniques of matched asymptotic expansions to second order in the mass ratio. The constraints from matched asymptotics at the subleading order in the matching equations then constrains the worldline in much the same way that the first order self force does. The broad description of the iterative matched asymptotics procedure belies the difficulty of the perturbative expansion. Notably, the singular field must now be determined through quadratic order in the mass, as must the singular part which arises from cross terms between the regular part and the singular part at first order.

In both derivations, the accelerated worldline is derived by insisting on a particular set of gauges for the interior and exterior metric expansions. Although the gauge choice differs between the derivations, the main feature that both derivations share is that they expand the metric in a gauge which is a 'fixed' gauge about the worldline for the interior expansion, and have a gauge condition similar to the Lorenz gauge condition in the exterior expansion.

The main qualitative difference between the derivation by Pound [42] and that by Gralla [43] are their respective treatments of the worldline. In the derivation by Pound [42], the worldline is

treated non-perturbatively. The resulting self force is the remarkably simple and elegant result that the full worldline follows a geodesic in the appropriately regularized spacetime,

$$\begin{aligned} a^{(2)\mu} = & -\frac{1}{2}(g^{\mu\nu}+u^\mu u^\nu)(g_\nu{}^\rho - h^{(1)R\,\rho}_\nu)(2h^{(1)R}_{\rho\sigma;\lambda} - h^{(1)R}_{\sigma\lambda;\rho})u^\sigma u^\rho \\ & -\frac{1}{2}(g^{\mu\nu}+u^\mu u^\nu)(2h^{(2)R}_{\nu\sigma;\lambda} - h^{(2)R}_{\sigma\lambda;\nu})u^\sigma u^\lambda, \end{aligned} \tag{1.35}$$

and in gauges smoothly related to the Lorenz gauge. In the derivation by Gralla [43], the worldline is treated perturbatively, so the expanded expression explicitly depends on the first order worldline. In a later publication which explored the subtle differences in the gauge freedom under the differing choices of worldline treatment, Pound confirmed [44] that if the non-perturbative worldline result is re-expanded with an appropriate construction of the gauge freedom in the worldline definition, the second order self force derived by Gralla is reproduced.

The treatment of the worldline as the sum of the background geodesic and a set of small corrections is potentially troublesome for use in a calculation of an extreme mass ratio inspiral. Over the course of the $\sim M/\varepsilon$ radiation-reaction time, the deviation from a background geodesic will no longer be perturbatively small, which would invalidate the assumption of the perturbative expansion. Instead, an inspiral would need to be 'sewed together' from several short-timescale expansions, which would lose the valuable sub-radian phase accuracy.

For the purposes of the multiscale treatment discussed in Chapters 4, 5, and 6, either worldline treatment could be incorporated. We choose in the work presented in this dissertation to base the perturbative framework more closely on the non-perturbative treatment of the worldline, and expand the worldline dependent expressions as necessary, via the formalism described in [44].

1.3.4 Strategies for self force evaluation and worldline evolution

The determination of the metric dependence of the equation of motion for the worldline from sections 1.3.2 and 1.3.3 is merely the starting point for an extreme mass ratio inspiral computation. In principal, the self force must be determined at each point along the forced orbit, and the orbit appropriately integrated to determine the full inspiral. The computation is not so direct, however, because the self force depends directly on the metric perturbation sourced by the small object, often at significant timelike separation. This echoed backreaction is known as the 'tail' part of

the self-field, and arises from the gravitational perturbation scattering off strong curvature of the large companion. The evaluation of the self force itself is a delicate matter, as often the full metric perturbation sourced by a pointlike body is the convenient quantity to simulate. The resulting metric is singular at the worldline of the small companion, so a numerical regularization scheme is required. An excellent review of these computational methods is given by [45], and I review here the basic features of the salient regularization methods.

The mode-sum approach [46, 47] relies on a series of analytically derived regularization parameters which are subtracted from the numerically determined spectral decomposition of the spacetime modes. Once the regularization parameters have been subtracted, the remaining numerical solution is smooth at the position of the worldline, and can be used to determine the self force.

The effective source method, first proposed by [48, 49], can be productively applied to address the singularity near the worldline, without specifying a specific mode decomposition. The main reasoning in the effective source method is that the problematic singular behavior arises from an analytically known 'puncture' metric. Therefore, the desired equation for the regular part of the metric perturbation may be written as

$$E_{\mu\nu}[h^{\mathcal{R}(1)}] = 8\pi T_{\mu\nu} - E_{\mu\nu}[h^{\mathcal{P}(1)}], \tag{1.36}$$

where $h^{\mathcal{R}(1)}$ is the residual, smooth field, and $h^{\mathcal{P}(1)}$ is the analytically determined puncture. The residual field can then be simulated directly and used to determine the self force.

The most recently applied method for regularization and self force evaluation is the worldline convolution method, which explicitly generates the Green's function bitensors in the computational domain. This method is the most closely related to the formal evaluation of the various components of the field based on the Green's function properties. Once the Green's function is computed over the computational domain, the evaluation of the self force at any point is the comparatively inexpensive task of integrating one of the Green's function arguments over the past worldline until the desired accuracy is attained.

Various strategies have been developed to efficiently compute or approximate the resulting self forced worldline evolution. In the remainder of this section, I review briefly the most commonly used methods in the self force community, and the costs and benefits of each.

The osculating geodesics method [28] takes advantage of the fact that, for extreme mass ratio inspirals, the instantaneous worldline will be very close to geodesic. Therefore, instead of evolving a full, forced worldline, the osculating geodesics method finds the self force for a several geodesic orbits, then uses that force information to evolve the orbit gradually through the space of geodesics. The osculating geodesic method loses subtle tail effects, as the self force at each step in the simulation is treated as though the object has been eternally on the current tangent geodesic. The error accumulated from this method should be similar to the scale of the second-order dissipative self force, and therefore gives rise to an $\mathcal{O}(1)$ phase error. The main efficiency gain from the osculating geodesic method is that a catalog of geodesic self forces can be computed and cached, so several different orbits can be evolved at various mass ratios using the same catalog of osculating geodesic self forces. A closely related method was recently presented [50], which also uses geodesic-based self forces to evolve the worldline, but takes advantage of near-identity transformations to achieve extremely rapid evolution of the worldline.

The self consistent evolution method [30] is considered the most accurate strategy yet implemented. Unfortunately, it is also by far the most costly method. Instead of developing some approximate representation for either the worldline or the metric perturbation, it evolves both the full worldline and the metric perturbation in lockstep. At each time interval of the simulation, the information from the forced worldline on the previous time step is back-reacted to update the metric perturbation, and simultaneously the metric perturbation from the previous step is used to determine the new position of the worldline. The self consistent successfully handles all of the radiation reaction effects for the slow evolution of the orbit. However, as I discuss in Section 4.4.2, it does not account for the slow evolution of the mass and spin of the central black hole during the inspiral, which can in principle result in $\mathcal{O}(1)$ phase errors over the inspiral. The self-consistent method can be modified to include the slow mass and spin evolution, as is described the appendix of Section 4.C.2.

1.3.5 | Waveform generation from EMRIs

Just as there is a tradeoff between speed and accuracy for the computation of the self forced worldline, there are various methods that may be used to evaluate the emitted waveform with

similar tradeoffs.

The most accurate and mathematically direct, but the most computationally expensive method is a direct extraction of the waveform from from metric perturbations sourced by the small companion [51]. In the two-step framework of first developing a self-forced worldline, then computing a waveform, this waveform generation technique ammounts to a full re-evaluation of the sourced Einstein field equation over the full inspiral. Some reuse may be made of the metric perturbations if they were already generated in the process of evolving the worldline, as in the self-consistent evolution, but direct evaluation of the waveforms remains very costly.

The technique of snapshot evolution of the waveform is more economical [52], and is analogous to the osculating geodesics method of worldline evolution. The waveform is first computed thoroughly for a catalog of geodesic orbits. Once such a catalog has been assembled, many self forced waveforms can be constructed by stitching together the different waveforms as the small companion gradually moves through the space of possible geodesics.

The most rapid method so far implemented, but also the most crude is the semi-relativistic approximation [53,54]. The semi-relativistic approximation derives an approximate time-dependent quadrupole source associated with the forced orbit, as though the motion was executed by point masses in flat space coordinates naively identified with the Boyer-Lindquist coordinates. The quadrupole is then used in the classic gravitational-wave quadrupolar source formula [26] to obtain an approximate waveform.

Chapter 1 Bibliography

[1] A. Einstein, *Uber Gravitationswellen*, Stizung of the Pruss. Acad. **February** (1918), pp. 154–167

[2] B. P. Abbott *et al.*, *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116** (2016) (6), p. 061102

[3] P. G. Bergmann, *Summary of the Chapel Hill Conference*, Rev. Mod. Phys. **29** (1957), pp. 352–354,
URL: https://link.aps.org/doi/10.1103/RevModPhys.29.352

[4] A. Einstein, *Approximative Integration of the Field Equations of Gravitation* (1916),
URL: https://einsteinpapers.press.princeton.edu/vol6-trans/213

[5] J. Weber, *Evidence for Discovery of Gravitational Radiation*, Phys. Rev. Lett. **22** (1969), pp. 1320–1324,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.22.1320

[6] R. A. Hulse and J. H. Taylor, *Discovery of a pulsar in a binary system*, Astrophys. J. **195** (1975), pp. L51–L53

[7] J. M. Weisberg and J. H. Taylor, *Relativistic binary pulsar B1913+16: Thirty years of observations and analysis*, ASP Conf. Ser. **328** (2005), p. 25

[8] S. Dwyer, *Advanced LIGO status*, J. Phys. Conf. Ser. **610** (2015) (1), p. 012013

[9] G. O. A. Team, *The ESA-L3 Gravitational Wave Mission* (2016)

[10] R. M. Wald, *General Relativity* (Chicago Univ. Pr., Chicago, USA, 1984)

[11] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973)

[12] S. Weinberg, *Gravitation and Cosmology* (John Wiley and Sons, New York, 1972), URL: http://www-spires.fnal.gov/spires/find/books/www?cl=QC6.W431

[13] A. Abramovici *et al.*, *LIGO: The Laser interferometer gravitational wave observatory*, Science **256** (1992), pp. 325–333

[14] S. Singh, L. A. De Lorenzo, I. Pikovski and K. C. Schwab, *Detecting continuous gravitational waves with superfluid ^{4}He*, New J. Phys. **19** (2017) (7), p. 073023

[15] F. Pretorius, *Evolution of binary black hole spacetimes*, Phys. Rev. Lett. **95** (2005), p. 121101

[16] L. Blanchet, *Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries*, Living Rev. Rel. **17** (2014), p. 2

[17] E. Racine and E. E. Flanagan, *Post-1-Newtonian equations of motion for systems of arbitrarily structured bodies*, Phys. Rev. **D71** (2005), p. 044010, [Erratum: Phys. Rev.D88,no.8,089903(2013)]

[18] A. G. Shah, J. L. Friedman and B. F. Whiting, *Finding high-order analytic post-Newtonian parameters from a high-precision numerical self-force calculation*, Phys. Rev. **D89** (2014) (6), p. 064042

[19] L. Blanchet, G. Faye and B. F. Whiting, *High-order half-integral conservative post-Newtonian coefficients in the redshift factor of black hole binaries*, Phys. Rev. **D90** (2014) (4), p. 044017

[20] T. W. Baumgarte and S. L. Shapiro, *Numerical Relativity: Solving Einstein's Equations on the Computer* (Cambridge University Press, 2010)

[21] F. Loffler *et al.*, *The Einstein Toolkit: A Community Computational Infrastructure for Relativistic Astrophysics*, Class. Quant. Grav. **29** (2012), p. 115001

[22] M. A. Scheel, M. Boyle, T. Chu, L. E. Kidder, K. D. Matthews and H. P. Pfeiffer, *High-accuracy waveforms for binary black hole inspiral, merger, and ringdown*, Phys. Rev. **D79** (2009), p. 024003

[23] R. Haas *et al.*, *Simulations of inspiraling and merging double neutron stars using the Spectral Einstein Code*, Phys. Rev. **D93** (2016) (12), p. 124062

[24] P. Grandclement and J. Novak, *Spectral methods for numerical relativity*, Living Rev. Rel. **12** (2009), p. 1

[25] F. HÃľbert, L. E. Kidder and S. A. Teukolsky, *General-relativistic neutron star evolutions with the discontinuous Galerkin method* (2018)

[26] S. W. Hawking and W. Israel (eds.), *THREE HUNDRED YEARS OF GRAVITATION* (1987)

[27] E. Poisson, A. Pound and I. Vega, *The Motion of point particles in curved spacetime*, Living Rev. Rel. **14** (2011), p. 7

[28] A. Pound and E. Poisson, *Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals*, Phys. Rev. **D77** (2008), p. 044013

[29] S. E. Gralla and R. M. Wald, *A Rigorous Derivation of Gravitational Self-force*, Class. Quant. Grav. **25** (2008), p. 205009, [Erratum: Class. Quant. Grav.28,159501(2011)]

[30] A. Pound, *Self-consistent gravitational self-force*, Phys. Rev. **D81** (2010), p. 024023

[31] H. A. Lorentz, *Theory of Electrons* (Dover, 2011, reprinted from 1915)

[32] M. Abraham, Ann. Physik **10** (1903), p. 105

[33] P. A. M. Dirac, *Classical Theory of Radiating Electrons*, Proceedings of the Royal Society of London A: Mathematical, Physic al and Engineering Sciences **167** (1938) (929), pp. 148–169

[34] S. E. Gralla, A. I. Harte and R. M. Wald, *A Rigorous Derivation of Electromagnetic Self-force*, Phys. Rev. **D80** (2009), p. 024031

[35] D. J. Griffiths, *Introduction to Electrodynamics* (Pearson, 1989)

[36] J. D. Jackson, *Classical Electrodynamics* (Wiley, 1998)

[37] Y. Mino, M. Sasaki and T. Tanaka, *Gravitational radiation reaction to a particle motion*, Phys. Rev. **D55** (1997), pp. 3457–3476

[38] T. C. Quinn and R. M. Wald, *An Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved space-time*, Phys. Rev. **D56** (1997), pp. 3381–3394

[39] A. Pound and J. Miller, *Practical, covariant puncture for second-order self-force calculations*, Phys. Rev. **D89** (2014) (10), p. 104020

[40] S. L. Detweiler and B. F. Whiting, *Selfforce via a Green's function decomposition*, Phys. Rev. **D67** (2003), p. 024025

[41] A. Pound, *Second-order gravitational self-force*, Phys. Rev. Lett. **109** (2012), p. 051101

[42] A. Pound, *Nonlinear gravitational self-force: second-order equation of motion*, Phys. Rev. **D95** (2017) (10), p. 104056

[43] S. E. Gralla, *Second Order Gravitational Self Force*, Phys. Rev. **D85** (2012), p. 124011

[44] A. Pound, *Gauge and motion in perturbation theory*, Phys. Rev. **D92** (2015) (4), p. 044021

[45] B. Wardell and A. Gopakumar, *Self-force: Computational Strategies*, Fund. Theor. Phys. **179** (2015), pp. 487–522

[46] L. Barack and A. Ori, *Mode sum regularization approach for the selfforce in black hole space-time*, Phys. Rev. **D61** (2000), p. 061502

[47] L. Barack, Y. Mino, H. Nakano, A. Ori and M. Sasaki, *Calculating the gravitational selfforce in Schwarzschild space-time*, Phys. Rev. Lett. **88** (2002), p. 091101

[48] L. Barack and D. A. Golbourn, *Scalar-field perturbations from a particle orbiting a black hole using numerical evolution in 2+1 dimensions*, Phys. Rev. **D76** (2007), p. 044020

[49] I. Vega and S. L. Detweiler, *Regularization of fields for self-force problems in curved spacetime: Foundations and a time-domain application*, Phys. Rev. **D77** (2008), p. 084008

[50] M. Van De Meent and N. Warburton, *Fast Self-forced Inspirals* (2018)

[51] N. Warburton, S. Akcay, L. Barack, J. R. Gair and N. Sago, *Evolution of inspiral orbits around a Schwarzschild black hole*, Phys. Rev. **D85** (2012), p. 061501

[52] S. A. Hughes, S. Drasco, E. E. Flanagan and J. Franklin, *Gravitational radiation reaction and inspiral waveforms in the adiabatic limit*, Phys. Rev. Lett. **94** (2005), p. 221101

[53] R. Ruffini and M. Sasaki, *On a Semirelativistic Treatment of the Gravitational Radiation From a Mass Thrusted Into a Black Hole*, Prog. Theor. Phys. **66** (1981), pp. 1627–1638

[54] S. Babak, H. Fang, J. R. Gair, K. Glampedakis and S. A. Hughes, *'Kludge' gravitational waveforms for a test-body orbiting a Kerr black hole*, Phys. Rev. **D75** (2007), p. 024005, [Erratum: Phys. Rev.D77,04990(2008)]

Chapter 2

Modern gravitational wave detectors

2.1 Gravitational wave detection strategies

2.1.1 The magnitude of the task

The original argument for the feasibilty of the detection of gravitational waves was based on broad computations of the magnitude of the most likely sources in the relevent frequency range: compact binaries. In particular, the source amplitudes can be estimated using the quadrupole formula (see Section 1.2.3) [55], as

$$h \sim \left(\frac{GM_{\text{eff}}}{rc^2}\right)\left(\frac{v}{c}\right)^2 \approx 10^{-16}\left(\frac{M_{\text{eff}}}{M_\odot}\right)\left(\frac{v}{c}\right)^2\left(\frac{1\text{kpc}}{r}\right), \tag{2.1}$$

where the effective mass M_{eff} is the mass needed for a pair of pointlike masses to produce an equivalent quadrupole moment to the approximated source. For instance, we should expect a characteristic strain of a pair of 30 $M_\odot$ black holes with a final velocity of $v \sim .1$ c at a distance of $r \sim 100$ Mpc to be of the order $h \sim 10^{-21}$, which is a reasonable estimate of the magnitude found by more detailed numerical simulations. For a more in-depth discussion of binary source modelling techniques, see Section 1.2.2.

The measurable effect of the gravitational wave passage (Section 1.2.1) is of a periodic variation of distances between free observers, or the periodic force of a bound resonator. Resonator experiments have been proposed and attempted, but none to date have been successful. Ground based gravitational wave detectors use Michelson interferometry to detect minute changes in the positions

of large, noise-isolated test-mass mirrors. While these mirrors are not strictly in free-fall, the pendulum suspension ensures that they react as though they were freely falling to periodic sources in the relevant frequency range ($\sim 1 - 1000$ Hz).

The relevant comparison for the purposes of determining the efficacy of a gravitational wave detector is between a relevant signal strength $h(\omega)$ and the noise of the detector. For visualization purposes, the signal to noise ratio can be approximated as [56],

$$\frac{S}{N} \approx \frac{h_c}{h_{rms}}, \tag{2.2}$$

Here, h_c is the characteristic strain, defined by $h_c = h(f)\sqrt{n}$, where n is the number of cycles that the source transmits during its lifetime. The root mean square (rms) noise is $h_{rms} = \sqrt{f_c}\tilde{h}(f_c)$ for a one-cycle-long signal at frequency f_c. In Fig. 2.1, several detector arrangements' noise characteristic strains $\tilde{h}(f)$ are plotted, along with characteristic source amplitude $h_c(f)$ for many interesting astrophysical sources. As the ratio between the characteristic strain of the source and the characteristic strain of the detector sensitivity is directly representative of the estimated signal to noise ratio, sources well above the sensitivity curves are easily detected, sources close above the curves may be detected, and sources below the curves are not detectable.

2.1.2 | Overview of detection strategy

Current and near-future gravitational wave detectors achieve the necessary extreme sensitivity in position measurement via laser interferometry. The basic optical setup for the use of Michelson interferometers for precision position measurement follows

1. A coherent (laser) source is directed to a beamsplitter, which diverts equal amplitudes of coherent waves down a pair of long paths (arms)

2. At the termination of the long arms, a pair of reflectors direct the waves back along the paths on which they've come. The effect is that once the pair of coherent waves return to the beamsplitter, they will acquire a relative phase associated with the sub-wavelength difference in the mirror positions.

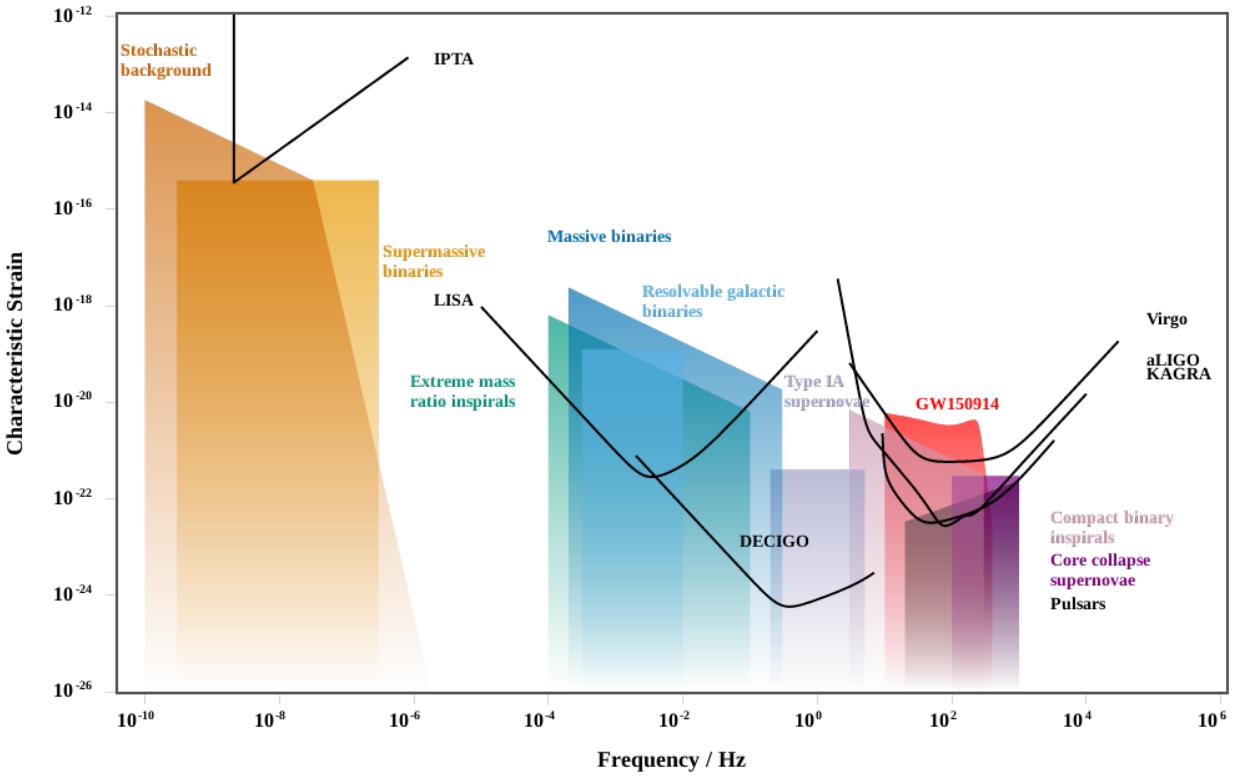

Figure 2.1: A plot of the noise curves of each of the current and proposed gravitational wave projects, against the anticipated spectrum of possible sources, created using tools authored by [57]. Plotted against various anticipated source strains are the characteristic strains as a function of frequencies, for both the current gravitational wave experiments (IPTA,Virgo, aLIGO), and for near-future planned experiments or experiments under constructions (LISA, DECIGO, KAGRA). The advanced Laser Interferometry Gravitational wave Observatory (aLIGO) is the american ground-based interferometry network, which made the first detections of gravitational waves. Virgo is the Italian ground-based intergerometer which particpated in the most recent detections with aLIGO. The International Pulsar Timing Array (IPTA) endeavors to detect very low-frequency gravitational waves associated with stochastic background or supermassive binaries by careful monitoring of the very regular radio pulses from high-quality radio pulsars. The Laser Interferometry Space Antenna (LISA) and the Deci-hertz Interferometry Gravitational wave Observatory (DECIGO) are future space-based gravitational wave experiments that will be formed of trios of satellites.

3. The two waves are recombined at the output port of the beamsplitter, which sums the coherent waves, resulting in an amplitude determined by interference. Under careful tuning, this output amplitude can be chosen to be precisely 0 (so that the two lengths of the interferometer legs differ by $n + 1/2$ wavelengths, for integer n).

4. Small fluctuations in the difference of distances to the mirrors now result in deviation from zero luminosity output from the beamsplitter.

Under ideal performance, a half-wavelength difference in mirror position will result in an output luminosity change similar to the original luminosity of the input laser. It is now easy to see that a high-power system could achieve sensitivity vastly finer than the wavelength of light used. This is the basic principle of the high-power laser interferometry systems used and proposed for gravitational wave observations.

The currently operational interferometers are the ground-based detectors, LIGO-Hanford, LIGO-Livingston, and Virgo in Italy. The LIGO detector network has used laser interferometry to achieve the first direct detection of gravitational waves, and to date has detected 5 binary black hole mergers as well as a celebrated multi-messenger binary neutron star merger. Terrestrial interferometers have sensitivity primarily in the frequency range $1 - 10^4$ Hz.

It is worthwhile to note the alternative method of gravitational wave observation used for extremely low wavelengths known as pulsar timing arrays. Millisecond radio pulsars act as reliable and fast clocks, due to the predictable nature of the signal from rapidly spinning neutron stars. Under a passage of gravitational waves, the individual fast pulses will acquire a small time delay, which would be detectable with radio telescopes. Currently, the pulsar timing array collaboration Nanograv has only published upper limits on the stochastic background, but it is hoped that with a data set including pulsar data through the year 2019 will provide a positive detection [58]. Pulsar timing arrays have sensitivities in the bands associated with the line-of-sight distance to the timing pulsars, so the range of detectable frequencies for Pulsar timing arrays is $10^{-10} - 10^{-7}$ Hz.

Space-based gravitational wave detectors, in particular the Laser Interferometer Space Antenna (LISA), will provide the ability to detect a vastly different frequency band from other detectors. The comparative noise-free environment of space provides an ideal context to perform the precision

distance measurements involved in gravitational wave observation. A space-based detector of the LISA design would have arm lengths determined by the separation of the constituent satellites, which will be $\sim 2.5 \cdot 10^9$ m. The frequency range of such a space-based gravitational wave detector would then be $10^{-4} - 10^{-1}$ Hz.

2.1.3 Matched filtering

Despite the incredible developments that have gone into reducing the noise in gravitational wave detectors [59], gravitational wave signals are very difficult to extract from the noise. If the LIGO collaboration did not know ahead of time what the signal of gravitational waves would be, the signal to noise ratio might be an insurmountable problem. However, thanks to many theoretical efforts, the gravitational wave signal for several interesting systems has been successfully computed. These computations allow the use of matched filtering techniques, which greatly enhance the signal to noise ratio and extract subtle gravitational wave data from the dominant noise. In this section, I review the basics of matched filtering techniques for contaminated signals.

2.1.3.1 Spectral noise density

A useful method of analyzing the the signal and noise characteristics of an measurement or transmission is to decompose the respective components into Fourier modes,

$$\tilde{h}(f) = \int_{-\infty}^{\infty} h(t) e^{2\pi i t f} dt. \tag{2.3}$$

The energy of a propagating wave varies as the square of the amplitude, so a physically interesting value is the spectral density, or power spectrum:

$$S_h(f) = \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} \left[h(t) - \bar{h}\right] e^{2\pi i f t} dt \right|^2, \tag{2.4}$$

where the normalization and the limit are used to ensure that the Fourier transform gives finite values. This is a measure of the amount of energy in the confounding noise sources between the frequencies f and $f + df$. In LIGO, the noise spectrum is bounded broadly by seismic noise at low frequencies, thermal noise of the mirrors at intermediate frequencies $40s^{-1} \leq f \leq 200s^{-1}$, and the photon Shot noise at high frequencies.

The LIGO spectral noise density has many sharp peaks in addition to its broad behavior, which are a result of resonance frequencies of the interferometer construction. At these narrow frequency bands, the detector is less able to detect gravitational waves on account of local contamination. However, these resonances are very narrow and do not substantially affect LIGO's detection rate. The LIGO sensitivity curve during the most recent observation run (O2) is depicted in Fig. 2.1.4.

2.1.3.2 Wiener optimal filter

For this section, I will follow the derivation route suggested in [60]. A general model for the LIGO detection scenario is that of some signal strength $h(t)$, and a noise function $\tilde{h}(t)$. We assume that through careful modeling, we know the function $h(t)$, and can use that knowledge to enhance the likelihood of detection. The overall measured value is then,

$$h_{tot}(t) = h(t) + \tilde{h}(t). \tag{2.5}$$

Define a filter for this measurement $K(t)$, so that the resulting overall signal and noise values are given by convolutions of the signal function and random noise variable with the filter function,

$$S = \int_{-\infty}^{\infty} K(t')h(t')dt', \tag{2.6a}$$

$$N(t) = \int_{-\infty}^{\infty} K(t-t')\tilde{h}(t')dt'. \tag{2.6b}$$

First, we wish to calculate $\langle N^2 \rangle$, which will be representative of the strength of the noise after the filter has been applied.

$$\langle N^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} dt \left(\int_{-\infty}^{\infty} K(t-t')\tilde{h}(t')dt' \right)^2 \tag{2.7}$$

By re-arranging the integrals, and re-expressing the arguments in frequency space, one finds the compact expression,

$$\langle N^2 \rangle = \int_{-\infty}^{\infty} df |\tilde{K}(f)|^2 S_h(f), \tag{2.8}$$

where the function S_h is the spectral noise density discussed in the previous Section, 2.1.3.1.

For this sort of difficult detection scenario, we wish to use our knowledge of what the signal should be to construct an optimal filter. For instance, if we were attempting to detect a sinusoid of a particular frequency, we could filter the measurement by a narrow band-pass filter around the

frequency we were looking for. The optimal filter for an arbitrary signal is the filter which maximizes the signal to noise ratio $S^2/\langle N^2\rangle$ for the detection of that signal. Similar to the manipulations above for the noise value, we can re-write the signal function in terms of the frequency spectrum,

$$\begin{aligned} S = \int_{-\infty}^{\infty} K(t)s(t)dt &= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} \tilde{K}(f)e^{-2\pi i f t}s(t)df \\ &= \int_{-\infty}^{\infty} df \tilde{K}(f)\tilde{s}^*(f). \end{aligned} \tag{2.9}$$

The optimal filter is computed by extremizing the signal-to-noise ratio with respect to the Fourier-transformed filter function $\tilde{K}(f)$. The optimal filter then satisfies,

$$\int \delta\tilde{K}(f)\tilde{s}^*(f) \int |\tilde{K}(f)|^2 S_h(f) = \int \tilde{K}(f)\tilde{s}^*(f) \int \tilde{K}(f)\delta\tilde{K}(f)S_h(f), \tag{2.10}$$

The condition (2.10) is satisified only if the filter function is precisely the signal to be extracted, $\tilde{K}(f) = \tilde{s}(f)$. Therefore, given the use of an optimal matched filter, the signal to noise ratio is,

$$\frac{S^2}{\langle N^2\rangle} = 4\int_0^{\infty} \frac{|\tilde{s}(f)|^2}{S_h(f)}df. \tag{2.11}$$

2.1.4 | LIGO/Virgo overview

The LIGO collaboration has the daunting task of detecting gravitational radiation, which interacts very weakly with matter. Through vast noise reduction efforts, the Advanced LIGO detectors are able to detect solar mass neutron star mergers to hundreds of Megaparsecs in distance. Even with the noise reduction efforts, a gravitational wave signal would be lost in the noise without the techniques of matched filtering, which rely on the prior knowledge of the signal functions that LIGO detects.

The spectral noise density for LIGO has a complicated shape, owing to the intricacies of the experimental setup and the resonances of the many components and vibration isolation systems. These considerations define the narrow peaks that can be seen in the spectral density plots. However, the most important limits of the noise density come from the broad low-lying shape of the noise curve. The broad shape of the curve is set by the seismic, thermal, and Shot noise at the low, mid, and high LIGO frequency ranges respectively. The resulting sensitivity curve as of the second Advanced LIGO observation run is depicted in Fig. 2.1.4, reproduced from [62].

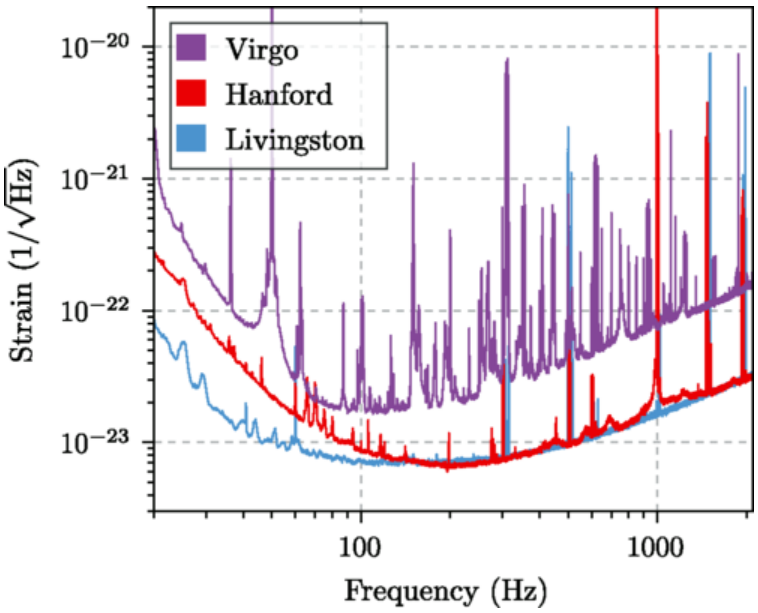

Figure 2.2: The LIGO/Virgo noise curves as of the second observing run (O2), which ended August of 2017 [61]. This plot shows the significan advances that have been made in the aLIGO noise curves, which are the low-lying blue and rend curves for the Livingston and Hanford detectors, respectively. The remaining curves show the rapid advancement of the Virgo detector towards design sensitivity. By late June, the detector was sufficiently sensitive to detect a binary neutrons tar inspiral to 10 Mpc, and in August, Virgo participated in the aLIGO/Virgo gravitational wave detections GW170814 and GW170817. This figure is reproduced from [62].

2.1.5 | LIGO/Virgo noise sources and mitigation

2.1.5.1 Seismic noise and mechanical isolation

The raw seismic root mean squared distance deviation as a function of frequency can be estimated as [63],

$$S_{\Delta x}^{1/2}(f) = 10^{-10}\left(\frac{100\ \text{s}^{-1}}{f}\right)^{3/2}\ \text{m s}^{-1/2} \tag{2.12}$$

This value is very high compared to the scale of the sources which Advanced LIGO seeks to observe, so the Advanced LIGO optical assembly makes use of thorough seismic isolation. The test masses are isolated by 3 vertical isolation stages, and 4 pendulum stages to provide a vibrationally quiet mirror for the gravitational wave measurements. Kip Thorne [63] introduces a quick method for estimating the attenuation from the seismic isolation systems: for each stage of isolation, the seismic vibration will be damped by a factor of $\sim (f_0/f)^2$, where f_0 represents the normal mode frequency

of the isolation system. From this simple estimate, which leads to a strain noise estimate of,

$$S_h^{1/2}(f) \approx 2.5 \cdot 10^{-14} \left(\frac{f_0}{f}\right)^{14} \left(\frac{100 \text{ s}^{-1}}{f}\right)^{3/2} \text{ s}^{-1/2}, \tag{2.13}$$

it becomes clear that the residual noise in the mirrors is exceptionally sensitive to the frequency of the isolation stages, and diminishes sharply with the frequency. This forms a crisp boundary at the typical frequency of the passive isolation system, which is softened when the active seismic isolation systems are taken into account. A detailed description of the active isolation can be found in LIGO releases [59].

2.1.5.2 Thermal noise and material choice

Broadly, one can estimate the magnitude of the thermal fluctuations by finding the deviation of the averaged position of the thermal vibrations of the molecules in the substance. The root-mean-squared vibration lengthscale is given by treating the material as thermally excited oscillators, giving a lengthscale $\sqrt{k_B T/m\omega^2}$. A typical order of magnitude estimate for a rigid material vibration frequency is $\omega \sim 10^{14}\,\text{s}^{-1}$ [63]. The broad beam of the LIGO laser averages the thermal oscillations of the material surface within its radius, and averages several oscillations of the individual material constituents. The area of the beamspot at the test mass mirrors is $1.2 \cdot 10^{-4}$ m^2, and on average, the beam spends $\sim 10^{-3}$ seconds in the Fabry-Perot arm of the interferometer.

For the remainder of this section, I review quantitative computations that have led to the LIGO noise curve estimates. The thermal noise in the interferometer setup manifests as the time-varying position of the mirror, as averaged over the beam power density, which obeys an approximately Gaussian distribution [60]:

$$q(t) = \int z(r, \phi; t) \frac{e^{-(r/r_0)^2}}{\pi r_0^2} r d\phi dr \tag{2.14}$$

Via the Generalized Fluctuation-Dissipation Theorem, [60] derives the spectral noise density of the thermal fluctuations of the test mass and support wires as,

$$S_{\Delta x}(f) = \frac{8W_{\text{diss}}}{F_0^2 2\pi f^2} \left(\frac{1}{2}hf + \frac{hf}{e^{hf/k_B T} - 1}\right) \approx \frac{8W_{\text{diss}} k_B T}{2\pi F_0^2 f^2}, \tag{2.15}$$

where the function W_{diss} is the dissipation power, and depends on various material properties of the test mass mirror. The careful materials analysis is given in [64], and extended to a finite test

Parameter	Symbol	Value [64]
Thermal expansion coeff.	α_l	$5.5 \cdot 10^{-7}$ K
Mass density	ρ	$2.2 \cdot 10^{3}$ kg/m^3
Specific Heat	C_v	$6.7 \cdot 10^{2}$ J/kg K
Poisson coeff.	σ	.17
Thermal Conductivity	λ^*	1.4 J/ m s K

Table 2.1: The relevant material properties of silicon dioxide for the computation of the thermal noise spectrum from the test mass

mass in [65]. The result of these computations is that, up to an error $\sim 10\%$ due to finite test mass corrections, the dissipation power is:

$$W_{\text{diss}} = \frac{(1+\sigma)^2 \lambda^* \alpha_l^2 T}{\sqrt{2\pi} C_v^2 \rho^2 r_0^3} F_0^2, \tag{2.16}$$

where σ is the Poisson coefficient, describing the expansion of a material orthogonal to an applied strain, λ^* is the thermal conductivity of the material, α_l is the thermal expansion coefficient, C_v is the specific heat, ρ is the density, and r_0 is the beamspot size. Using the the relevant quantities for fused silica [64] (given in Table 2.1), the broad estimate for the thermal noise spectrum becomes,

$$S_h(f)^{1/2} \approx (8.20) \cdot 10^{-24} \text{ s}^{-1/2} \left(\frac{1 \text{ s}^{-1}}{f} \right) \tag{2.17}$$

The thermal noise in the bulk fused silica is a real concern, but due to a wide laser spot and the choice of materials, it is largely mitigated relative to other sources of noise. The next possible source of substantial thermal noise is the coating on the test mass that provides the high reflection coefficient. This coating is made from tantalum pentoxide, a dielectric surface. The dominant contribution to this is the Brownian noise [66]. The authors of [67] quote the thermal noise spectrum for a coating as a function of the loss angle ϕ:

$$S_{\Delta x} = \frac{2\sqrt{2} k_B T \phi (1-\sigma^2)}{\pi^{3/2} f r_0 Y}, \tag{2.18}$$

where Y is the young's modulus of the material. The precise loss angles of this intricate setup are measured in [67].

2.1.5.3 Shot noise and laser power

The photon Shot noise arises from the discrepancy from a steady signal on account of the fact that photons are emitted discretely, rather than as a perfectly continuous wave. Instead, even very steady sources like the LIGO laser emit a superposition of many small wavepackets, each with energy $\hbar\omega$. Kip Thorne nicely reviews the Shot noise as well as many of the other topics covered in this brief review [60]. The spectrum of shot noise for the direct light power measurement is white:

$$S_w(f) = W\hbar\omega, \tag{2.19}$$

where W is the power of the light source. For examining the effect this has on LIGO detections, we must consider the additional filtering performed on the light before it is measured. Lyons et. al. derive the filter function between the spectral noise density of the position of the test mass and the output spectral noise density [68], $S_w(f) = |H(f)|^2 S_{\Delta x}(f)$. The filter function $H(f)$ depends greatly on the details of the LIGO interferometer construction, and is derived as [68],

$$|H(f)|^2 = 4\frac{\omega^2}{c^2} W_2 \frac{T_3^2 r_4^2}{(1 - r_3 r_4)^4} \frac{1}{(1 + (2\pi f/\omega_c)^2)}, \tag{2.20}$$

where W_2 is the laser power at the interferometer input, T_3 is the transmission coefficient of the partially reflecting Fabry Perot mirror at the start of the interferometer arms, and r_3 and r_4 are the reflection coefficients of the same Fabry Perot mirrors and the test mass mirrors, respectively. Finally, ω_c is the characteristic frequency of the Fabry Perot interferometer setup. All of the values for these characteristic parameters for the Advanced LIGO detectors are given in Table 2.2

To find the amount of displacement that the photon Shot noise would imitate in the LIGO arms, we compute the Shot noise arriving at the detector multiplied by the inverse of the filter function.

$$S_{\Delta x}(f) = \frac{S_w}{|H(f)|^2} \tag{2.21}$$

Working out all of the numerical factors from Table 2.2 gives an estimate for the power spectrum of the distance displacements,

$$S_{\Delta x}(f) = 2.60 \cdot 10^{-37} \left(1 + \left(\frac{2\pi f}{709\ \mathrm{s}^{-1}}\right)^2\right) \mathrm{m}^2\ \mathrm{s}^{-1}. \tag{2.22}$$

Parameter	Symbol	Value [59]
Power in interferometer arms	W_{arms}	750 kW
Laser power at interferometer input	W_2	5.2 kW
Absorbed power in test mass	W_{diss}	375 mW
Reflection coefficient of end test mass	r_4	$1 - 5 \cdot 10^{-6}$
Reflection coefficient of intermediate test mass	r_3	.986
Transmission coefficient of intermediate test mass	T_3	.014
Laser frequency	ω	$1.78 \cdot 10^{15}\ s^{-1}$
Cavity pole for Fabry Perot	ω_c	$(c/2l)(1 - r_3 r_4)/r_3 r_4 \approx 709\, s^{-1}$
Beamspot size	r_0	$6.2 \cdot 10^{-2}$ m

Table 2.2: Several important parameters for the Advanced LIGO interferometer, used in the text for estimating the effective filter function for the Shot noise

As a final step, we must take the square root and multiply by the LIGO arm length to convert this to a strain amplitude, as is traditionally plotted in the LIGO noise spectra,

$$S_h(f) = 1.53 \cdot 10^{-23} \left(1 + \left(\frac{2\pi f}{709 s^{-1}}\right)^2\right)^{1/2} \text{s}^{-1/2}. \tag{2.23}$$

2.1.6 KAGRA (future detector)

A new ground based gravitational wave detector, KAGRA, is under construction in the Kamioka mine in Japan [69]. KAGRA will be a novel addition to the detector network, as it will be the first gravitational wave observatory to use cryogenically cooled mirrors. The design constraint for cryogenically cooled components places limits on the other noise-mitigation systems. Most importantly, the maximum laser power in the Fabry-Perot cavities for KAGRA will be on the order of 10^{-3} of the LIGO laser power [70]. The lower laser power and thermal noise level has the practical upshot that KAGRA will be quantum limited by Shot noise over most of its spectrum. Recent work indicates that careful tuning of the various parameters of the detector system can partially mitigate some of non-thermal noise costs [71], with the anticipated sight range for binary neutron star events of approximately 150 Mpc. In addition to the usefulness of another high-sensitivity gravitational wave detector in a region very helpful for further sky localization, the exploration of cryogenically cooled components will provide valuable information on the possible design decisions when looking to next-generation detectors.

2.1.7 LIGO India (future detector)

The inclusion of an additional ground-based gravitational wave detector in the LIGO/Virgo interferometry network has been proposed [72] for construction in India. The Indian Initiative in Gravitational-Wave Observations submitted a proposal to the Indian government granting bodies in 2011, and secured an "In-Principle Approval for LIGO-India" in February 2016. The LIGO-India design proposal is identical to the Advanced LIGO design which has shown great success in detecting binary black-hole and binary neutron star mergers.

The addition of an Indian detector to the network would offer a number of important benefits. The most marked benefit will be the improved sky localization of many sources which LIGO can poorly resolve due to the comparatively similar detector orientations and positions of the American detectors [73]. The inclusion of LIGO-India will also boost the sensitivity of the network to gravitational wave events and provide additional coincidence verification to improve confidence intervals in detections. Finally, more orientations of detectors allows better constraints on the permitted polarizations of the incident gravitational waves, which is an important test of modified gravity theories.

2.1.8 Future space-based detectors

2.1.8.1 LISA

The next, and highly anticipated, extension of the gravitational wave detector spectra is the Laser Interferometry Space Antenna (LISA), which is a project led by ESA, with NASA participation [75]. The detector will be constructed as a triangular array of satellites, each of which equipped with a full, symmetric interferometry system. A significant advantage of the symmetric design is that two independent measurements of a gravitational perturbation can be made simultaneously, unlike a typical Michelson interferometer. The gain in sensitivity is approximately that which would occur from placing two similar Michelson systems at the same location. The proposal, now accepted by ESA, is for a symmetric detector of arm length $2.5 \cdot 10^6$ km, in an earth-trailing heliocentric orbit [76]. The full detector configuration is scheduled to launch in 2034.

One of the key requirements of the LISA project is sufficient control of the test masses. To

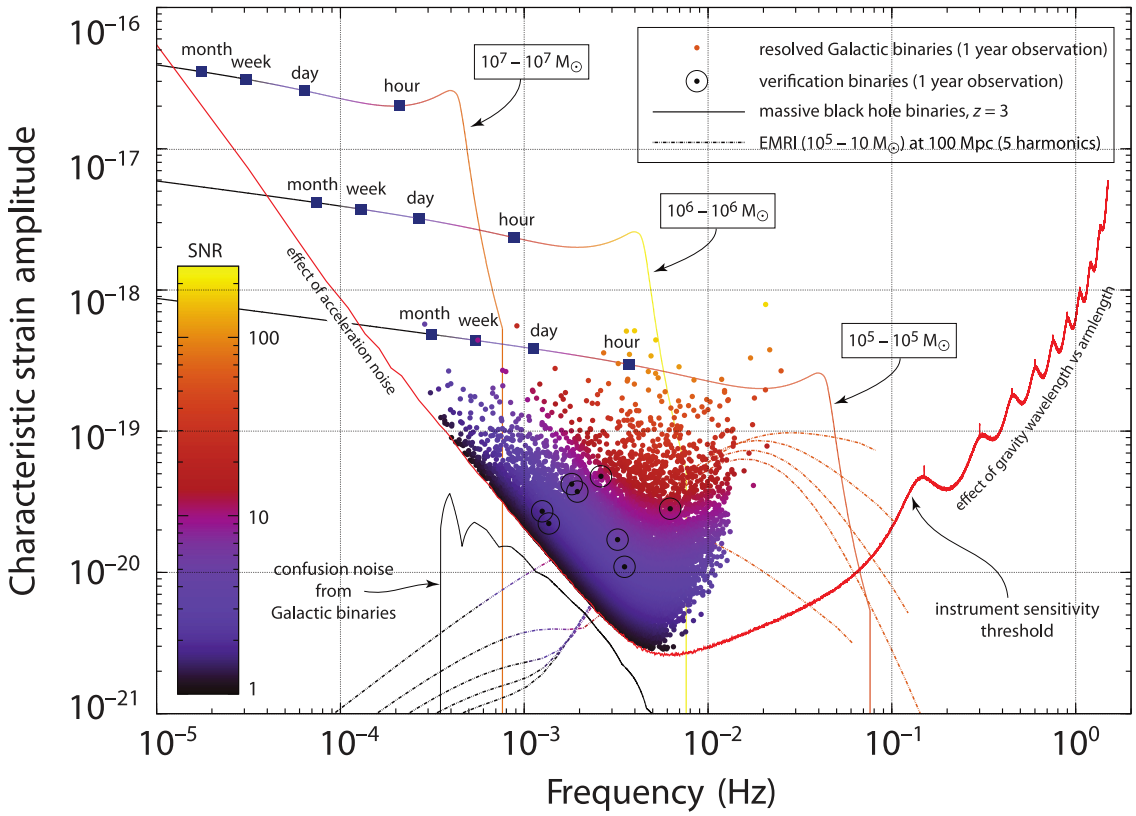

Figure 2.3: A plot of the characteristic strain amplitude vs frequency for the anticipated LISA noise curve, compared with several of the salient source magnitudes (for a discussion of characteristic strain, see 2.1.1). The loudest soruces will be supermassive black holebinaries, for which the evolution will sweep through the detector band over the course of the evolution. Also plotted are several resolvable galactic binaries, which will not significantly change in frequency over the course of the LISA mission. Finally, plotted low in the basin of the noise curve are several harmonics of an extreme mass ratio inspiral (EMRI), which will be well within the threshold of detection provided adequate source modelling. This plot was released in [74].

reliably detect gravitational perturbations, the test masses of the detector must be in steady freefall to sufficient precision. The LISA pathfinder project, which ran for years 2015-2017, was designed to test the technologies for the test mass control systems. For space-based detectors, this involves releasing the test mass to float freely within the craft and carefully controlling additional perturbations of the test masses from additional noise sources, such as electrostatic charging, residual accelerations, and thermal sources. Further, the craft surrounding the test masses is carefully controlled with femto-newton charged fluid thrusters [77]. The LISA pathfinder project completed in 2017, reporting great success. The February 2017 run outperformed the the LISA test mass design

requirements by a factor of 2 over much of the LISA spectrum [78].

The advent of space based gravitational wave instrumentation is highly anticipated due to the multitude and variety of sources that will be above the detection threshold. Observable sources will include numerous $(10^3 - 10^5)$ compact white dwarf binaries, compact neutron star and black hole binaries (similar to LIGO sources), massive black hole binaries, and, most interesting for this dissertation, extreme mass ratio inspirals (EMRIs) [74]. The ability to resolve sources which eventually become LIGO sources early in their inspiral is an exciting suggestion which may allow extremely precise measurements of particularly broad-spectrum events. The extreme mass ratio inspirals are events in which a compact object inspirals into a massive black hole, so have mass ratios $\sim 10^5 - 10^9$. Extreme mass ratio inspirals present an exciting prospect for future scientific results (see Section 2.3), but also a challenge for source modelling. The source modelling challenge is described in Section 1.3, and is the focus of the research developments in Chapters 4, 5, and 6. The anticipated noise curve of the LISA mission along with the characteristic strain of several important sources, from [74], is shown in Fig. 2.3.

2.1.8.2 DECIGO

The second planned space-based gravitational wave detector is the Japanese Deci-hertz Interferometer Gravitational wave Observatory (DECIGO). The observation goal for DECIGO is to fill the 'gap' in the frequency spectrum between the very low frequencies of LISA and the higher frequencies of LIGO. The project was first proposed to measure distant binary neutron star inspirals [79] and to further constrain the Hubble constant via the measurements of standard sirens, as has now been done using the first LIGO binary neutron star detection (discussed in Section 2.2.2.3).

The final design for DECIGO is a 1000 km arm-length detector, and has sensitivity primarily in the band $10^{-1} - 10$ Hz [80]. The full DECIGO project is planned for launch in the mid 2030s. Prior to launch of the full DECIGO design, a minimal pathfinder project for the demonstration of key technologies is hoped to launch in the 2020s [81]. The pathfinder project is called B-DECIGO (for "Basic" DECIGO), and will have a 100 km arm length. Optimistically, if the design goals are met, and the noise curve is primarily Shot-noise limited, B-DECIGO should have a sight range for loud binary black hole events events to redshift $z = 10$ [81].

2.2 Observed gravitational wave events

It is a tremendously exciting time for gravitational wave science: over the past three years, LIGO/Virgo has detected a total of five gravitational wave events positively identified as binary black hole events. Each of these events were sourced by the merger of black hole binaries with tens of solar masses, which grants new insight into a previously unobserved range of black hole masses and allowed qualitatively new tests of strong-field general relativity. Even more striking was the binary neutron star event, GW170817, which was detected first via gravitational wave signal in LIGO/Virgo. The trio of gravitational wave detectors was then able to assist electromagnetic telescopy to follow up on the event, which then provided the first trove of multi-messenger astronomy data from a strong-field event.

In this introductory section, I will discuss some of the most salient scientific conclusions from each of the type of events. For each event category, I will also discuss the additional information that might be gleaned with wide, numerous samples of such events, which will become available with more sensitive and more populated detector networks.

2.2.1 Binary black hole events

2.2.1.1 Overview

Five binary black hole systems have been observed by the LIGO/Virgo collaborations. For easy reference, I have collected the important details from each of the binaries into a compact table. The information for the table is extracted from the detection papers [62, 82–85].

The events detected by LIGO are extremely powerful events. During the binary merger of black holes, each of which possessing tens of solar masses, $\sim 10^{47}$J are radiated to gravitational waves, with a peak power on the order of $\sim 10^{49}$W. For many of the detected events, LIGO confirmed that over a full solar mass was converted to gravitational waves. This extreme power output is notable; it exceeds the total power of all visible electromagnetic sources. Due to the weakness of gravity and the distance to the source, the luminosity of the gravitational wave on the surface of the Earth is $\sim 1.5\text{mW/m}^2$, with a characteristic strain $\sim 10^{-21}$ [86]. LIGO, therefore, detects the periodic

Event	GW150914	GW151226	GW170104	GW170608	GW170814	GW170817
Signal-to-noise ratio ρ (filtered)	23.7	13.0	13	13	15	32.4
Primary mass $m_1/M_\odot$	$36.2^{+5.2}_{-3.8}$	$14.2^{+8.3}_{-3.7}$	$31.2^{+8.4}_{-6.0}$	12^{+7}_{-2}	$30.5^{+5.7}_{-3.0}$	1.36-1.60
Secondary mass $m_2/M_\odot$	$29.1^{+3.7}_{-4.4}$	$7.5^{+2.3}_{-2.3}$	$19.4^{+5.3}_{-5.9}$	7^{+2}_{-2}	$25.3^{+2.8}_{-4.2}$	1.17-1.36
Final mass $M_f/M_\odot$	$62.3^{+3.7}_{-3.1}$	$20.8^{+6.1}_{-1.7}$	$48.7^{+5.7}_{-4.6}$	$18.0^{+4.8}_{-0.9}$	$53.2^{+3.2}_{-2.5}$	$2.74^{+0.04}_{-0.01}$
Final spin a_f	$0.68^{+0.05}_{-0.06}$	$0.74^{+0.06}_{-0.06}$	$0.64^{+0.09}_{-0.20}$	$0.69^{+0.04}_{-0.05}$	$0.70^{+0.07}_{-0.05}$	
Radiated energy $E_{\rm rad}/(M_\odot c^2)$	$3.0^{+0.5}_{-0.4}$	$1.0^{+0.1}_{-0.2}$	$2.0^{+0.6}_{-0.7}$	$0.85^{+0.07}_{-0.17}$	$2.7^{+0.4}_{-0.3}$	$> .025$
Luminosity distance D_L/Mpc	420^{+150}_{-180}	440^{+180}_{-190}	880^{+450}_{-390}	340^{+140}_{-140}	540^{+130}_{-210}	40^{+8}_{-14}
Source redshift z	$0.09^{+0.03}_{-0.04}$	$0.09^{+0.03}_{-0.04}$	$0.18^{+0.08}_{-0.07}$	$0.07^{+0.03}_{-0.03}$	$0.11^{+0.03}_{-0.04}$	
Sky localization $\Delta\Omega/deg^2$	230	850	1200	860	60	28

Table 2.3: Summary table of the six identified gravitational wave events from LIGO/Virgo. This table comprises published data from [62,82–85]. Each of the events is named by its date of detection, and all of the first five columns are detections from binary black hole inspirals. The most recent event GW170817 is a detection of a binary neutron star inspiral, but specifically contains the information gleaned from the gravitational wave signal alone. The significant improvement in sky localization for the last two events are due to the participation of Virgo in late 2017.

motion of the test mass mirrors to within $\sim 10^{-18}$ m.

The number of distinct parameters extracted by LIGO (c.f. Fig. 2.3) and the precision at which they are measured is striking for what is effectively a one-dimensional data set (strain as a function of time). The reason for this comparative wealth of information is the richness of the gravitational wave signal itself (Fig 2.4). I will discuss the rough intuition for how these various parameters can be extracted from the waveform. In practice, a multidimensional fit is performed using credible models of the possible waveforms, but the rough reasoning below gives some intuition for why those fits work as well as they do. The waveform is frequently described as possessing three distinct dynamical regimes [88]. The long inspiral is well-described by post-Newtonian theory (high separation), the high-field interaction phase right before merger must be computed via numerical relativity, and the ringdown phase can be approximated by black hole perturbation theory.

The chirp mass of the binary $(m_1 m_2)^{3/5}/(m_1+m_2)^{1/5}$ is the parameter which determines the frequency scale of the inspiral phase of the merger. However, near the final stages, merger and ringdown, the mode structure of the Schwarzschild spacetime become most important, which depends most closely on the remaining total mass $M = m_1 + m_2$ (aside from radiation that has already escaped) [84]. The evolution of the dominant frequency over the course of the inspiral grants dis-

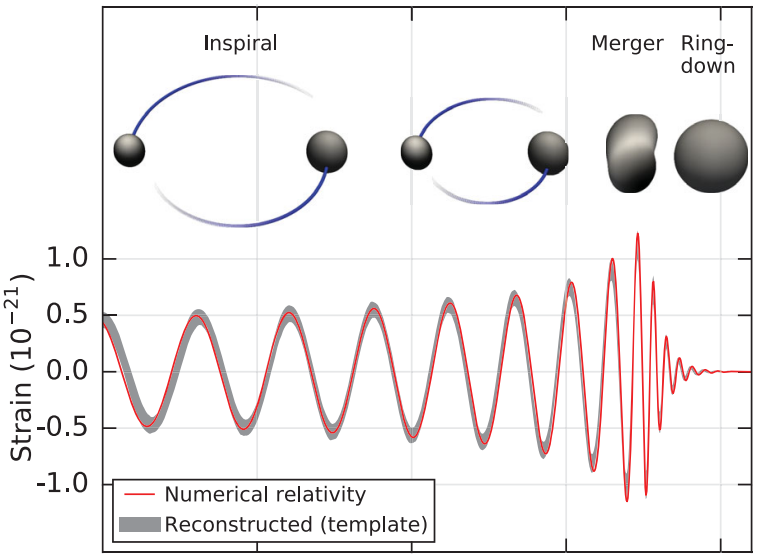

Figure 2.4: The best-fit waveform for the first detected gravitational wave in LIGO [87], along with the salient stages of the waveform. During the long, slow inspiral (much to the left of the figure), Post-Newtonian theory is relevant. During the final few orbits of the inspiral and merger, numerical relativity simulation is required. Finally, during ringdown, black hole perturbation theory can be used to compute the gravitational effects, but numerical relativity is often used to generate the waveform through ringdown.

criminating data sufficient to determine the initial masses and final mass. Therefore, comparison with numerical relativity [89,90] grants the additional information of the amount of energy radiated independent of, but consistent with, a 'balance law' total conservation of energy.

In addition to the binary parameters, general relativity computations predict a specific quantity of energy emission associated with a waveform shape. In astrophysics, similar useful sources are referred to as 'standard candles' and the gravitational analog has come to be known as 'standard sirens', as they encode distance information directly in the luminosity and the strain amplitude, respectively. The gravitational wave amplitude at the earth and the $1/r$ scale of the strain may be used to infer the luminosity distance. For many of the events detected so far, those distance estimates are fairly broad. With more precise measurement of the binary parameters, the luminosity distance inferred from the amplitude will also sharpen.

2.2.1.2 Population inference

The population of black holes in the universe is notoriously difficult to estimate. Modelling of stellar evolution still has many parameters which are not well constrained by modern observations, including the dynamics of gas envelopes in the viscinity of the formation of compact objects. A large collection of dynamical information must be included at high precision in order to develop an understanding of the number and mass spectrum of black holes [91] in absense of direct detections.

While estimates of the number and mass range of emitting stellar objects are available from all-sky surveys (e.g. [92]), these do not give direct information about black hole population unless the rates of collapse and merger are known. To make matters worse, there could be a remnant population of black holes with a distinct mass spectrum from primordial collapse [93]. These primordial black holes would also dynamically merge to generate a potentially complicated population of black holes at low redshift [94].

Observation of black hole systems can then start to be very relevant for astrophysical information about the outcomes of the various dynamical processes which might produce black holes. A population of low-mass black holes was previously cataloged by x-ray studies which are able to detect the emission from a black hole accretion disc, but LIGO has discovered an entirely different population of $\sim 10\ M_\odot$ black holes.

The current understanding of black hole populations is significantly improved by the new information from gravitational wave detectors. In the observation paper [85], the LIGO collaboration estimated the merger rate of black holes binaries such that the two companions each have $m_1, m_2 \geq 5M_\odot$ and final mass $M \leq 100M_\odot$ given the 4 merger events at the time of publication. The result of their statistical analysis is a rate $12 - 213\ \mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ at 90% confidence level. The range of the event rate inferred from LIGO during these observation runs may seem very broad, but it is a marked improvement over the ranges of rate estimates that were available only through detailed calculations prior to detection [95–97].

2.2.1.3 Constraints on theories of gravity

The gravitational wave events detected by the LIGO/Virgo collaboration are the first tests which truly probe the strong-field dynamics of general relativity. A number of theories of modified gravity have been suggested as possible alternatives to the particle physics construction of dark matter or dark energy sourced by unknown matter fields. The LIGO detections now permit comparisons between observed strong field dynamics, and those that would be predicted from modified gravity theories.

Many of the suggested alternative theories of gravity can be written as a coupled scalar-tensor theory in which a scalar field couples only via the metric contributions required to maintain Lorentz invariance of the action. Additionally, there is motivation from string theory to consider the possibility of axions (pseudoscalars) [98]. Other alternative metric theories of gravity include those with additional Lorentz-invariant contributions from the Riemann tensor ($R_{\mu\nu}R^{\mu\nu}$, $R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho}$, etc.), or $f(R)$ theories. Finally, large additional dimensions have been suggested as a possible cure to hierarchy problems, which is a theoretical framework known as the Randall-Sundrum model [99].

Aside from performing a full numerical relativity simulation of the strong-field dynamics in a modified theory of gravity, constraints can be placed on the more mild modifications of gravitational dynamics in the alternative theories. A frequently used method for testing the high-separation (and therefore low-speed) dynamics of an inspiral is known as the Parameterized post-Newtonian (PPN) method. Instead of performing a direct post-Newtonian expansion (discussed in Section 1.2.2.1), one performs an expansion with a number of free parameters which quantify the amount of curvature produced by rest masses, how much nonlinearity arises from gravitational perturbations, and several other important features of modified theories of gravity. PPN forms a useful phenomenological intermediary between the multitude of theories of modified gravity and the corresponding experimental tests. Instead of considering the details of modified theories, tests can constrain the PPN constants and new theory modifications can compare their predictions to the bounds.

The set of LIGO detections have started to place increasingly stringent constraints on the Post-Newtonian parameters for the merger dynamics [84]. All detections to date are comfortably compatible with the general relativity predictions (GR parameters are within the 90% confidence interval).

Also, due to the fidelity of the waveform to general relativity predictions, constraints may be placed on the vacuum dispersion relation of gravitational waves, which correspond to a failure of gravitational waves to consistently travel at the speed of light in vacuum. Multi-messenger astronomy (discussed below) is considerably more sensitive to this parameter, but binary black hole studies do offer some information.

In addition, there is the more subtle and more daunting constraint implied by the spectacular confirmation of the characteristic waveform associated with a gravitational wave 'chirp' from a binary merger in general relativity. By comparison, the Hulse-Taylor pulsar system granted significant constraints on the energy radiation, and any modified theory of gravity should be required to reproduce that energy loss rate within appropriate errors. With the advent of frequent direct gravitational wave detections, alternative theories of gravity should now be required to reproduce the full waveform to within the precision measured by LIGO.

Many alternative theories of gravity have not yet been shown to possess a well-posed Cauchy initial value problem [100,101]. For those that do have a well-posed Cauchy initial value problem, or can be treated perturbatively, there are relatively few simulations of the merger dynamics. Recent work has started to address this deficit for scalar-tensor gravity [102], and further careful simulation of proposed theories may reveal subtle dynamical consequences that constrain the deviation from general relativity.

2.2.1.4 Future tests from more numerous events

The projected improvements in the Advanced LIGO noise curve will extend the LIGO detection range by a factor of 2-3 over current performance [103]. With the additional detection volume, the statistics of the black hole binary merger rate will significantly improve (by the cube of the detection range). With improved statistics for the merger rate, the inferred population of black holes will become a resource for determining uncertain features in the computations initially used to estimate the black hole merger rate from astrophysical computations.

The rate of black hole merger is sensitive to the production mechanism and to the metallicity of the region in which the black holes are produced [95]. With high numbers of binary black hole observations from LIGO, the parameters associated with these formation scenarios might be

constrained. Further LIGO observations may also be able to more tightly constrain the population of primordial black holes, and the population of black holes that arise from cascaded mergers of primordial black holes. It remains an open question as to what portion of the dark matter budget might be explained by primordial black holes observable by LIGO [104].

Finally, it has been proposed that the population spectrum of black hole masses and spins could be suggestive of beyond the standard model physics. In particular, there is the ambitious suggestion that if a new field tended to form a bound state around a spinning black hole, it could undergo superradient scattering to produce a 'black hole bomb' [105, 106] scenario. Such dynamical effects within a Kerr ergosphere would rapidly rob spinning black holes of their angular momentum, leaving a population biased to near-Schwarzschild black holes. Current population statistics are insufficient to give significant information on such scenarios, but the suggestion offers a strong motivating example for the reasoning that can be performed with statistics from a large LIGO/Virgo detection volume.

2.2.2 Binary neutron star events

2.2.2.1 Overview

On August 17th of 2017, the LIGO and Virgo detectors observed for the first time the gravitational waves emitted from a binary neutron star merger [83]. The physical processes involved binary neutron star collisions run the full gamut of physically extreme systems, including strong-field relativistic effects, extreme electromagnetic fields to the order of $10^5 - 10^{11}$ T [107, 108] (for some neutron stars), and novel, unknown behavior of the strong nuclear interactions deep within. The August 2017 binary neutron star collision is the first event to leverage multi-messenger astronomy, using both gravitational and electromagnetic signals [109–116].

Neutron stars are extremely dense astrophysical objects, supported by the Fermi degeneracy pressure of nuclear matter. The high densities and pressures of neutron star interiors are still poorly understood, primarily owing to the complexity of the quantum chromodynamic interactions which become important for dense nuclear matter. Several models have been proposed for the interior states of neutron stars [117, 118], each of which can be formulated as a prediction for the neutron star equation of state [119].

The event GW170817 is remarkable for gravitational wave astronomy for a number of reasons. The detection and trigger announcement were sufficiently timely, and had a sufficiently precise sky localization for electromagnetic telescopes to follow up and locate the electromagnetic counterpart. Further, the Fermi gamma-ray telescope (which is largely direction independent, so did not need reorientation) was able to detect the near immediate accompanying gamma ray burst. This first publication of the multi-messenger gravitational and electromagnetic observations brought together 60 observational collaborations with over 2000 contributors [120].

2.2.2.2 Nuclear physics results

The disruption of the dense nuclear matter of binary neutron stars during collision gives rise to a host of nuclear physics interactions. The two primary pieces of information that can be extracted from the event are the gravitational wave data regarding the parameters of the merger and the time-dependent electromagnetic spectrum from the event regarding the nuclear content of the ejecta.

The gravitational wave signal sourced by a binary neutron star merger depends on the mass profile of the two neutron stars as well as the extent to which the neutron stars are distorted by the strong tidal fields during the merger, quantified as the tidal deformability [121, 122]. The tidal deformability affects the phase of the waveform. However, due to a degeneracy between aligned component of the spin and the mass ratio of the neutron stars, the LIGO/Virgo upper bound on the deformability depends weakly on the assumptions of the spin. The LIGO/Virgo collaborations constrain the combined deformability parameter,

$$\tilde{\Lambda} = \frac{16}{13}\frac{(m_1 + 12m_2)m_1^4\Lambda_1 + (m_2 + 12m_1)m_2^4\Lambda_2}{(m_1 + m_2)^5}, \tag{2.24}$$

where m_1 and m_2 are the primary and secondary neutron star masses, and Λ_1 and Λ_2 are the tidal deformabilities of the individual neutron stars, such that the quadrupole is related to the electric part of the Riemann tensor by

$$Q_{ij} = -\Lambda m^5 \mathcal{E}_{ij}. \tag{2.25}$$

The LIGO/Virgo collaborations find a bound of $\tilde{\Lambda} \leq 800$ at 90% confidence level. Typical ranges for neutron stars are $\sim 50 - 2000$ [121], and as the signal-to-noise ratio for these events improves, we can expect a direct measurement of the deformability. Notably, for future events, the differences in the

deformability across events is a more physically motivated quantity, due to the inherent ambiguity in relativistic definitions of the quadrupole [123].

Many of the heavier elements of the periodic chart, above iron, cannot be formed by the gradual synthesis that occurs during the nuclear fusion chain in stars. Instead, many elements require the rapid neutron-capture process, known as the 'r-process', to form in significant quantities [124]. The nuclear interaction requires a highly neutron-rich environment to increase the atomic mass number to form the heavier stable elements observed in the universe [125,126]. Supernovae and neutron star mergers have recently gained popularity as the strongest candidates for r-process synthesis [127,128]. The single binary neutron star multi-messenger observation was sufficient to lend great support of the hypothesis of neutron-neutron star as the dominant method of production of r-process elements.

Following merger, there are only a few possibilities for the resulting core object. It may form a new, more massive neutron star, which could potentially be stable depending on the neutron star equation of state and the maximum mass before gravitational collapse produces a black hole. Alternatively, if that maximum mass is exceeded by the final product of the merger, it should immediately undergo gravitational collapse to form a Kerr black hole. As a middle-ground option, there exists the possibility of forming an unstable hypermassive neutron star or supramassive neutron star, temporarily supported by extreme spin [129]. Such an over-massive neutron star will continue to lose energy and transition to the gravitational collapse scenario, but after a delay.

The outcome of these remnant studies further constrains the equation of state, which determines the maximum neutron star mass. In the LIGO/Virgo search for the gravitational wave emission from a possible hypermassive or supramassive neutron star, no gravitational waves were detected, but the upper limits are significantly greater than the predicted strain [130]. There is hope, then, for future detection of these remnant signals from detectors with improved sensitivity, including the upgraded LIGO detectors or next-generation detectors.

2.2.2.3 Constraints for theories of gravity

The most immediate constraint on alternative theories of gravity from the multi-messenger event is a test on the speed of propagation of gravitational radiation as compared to the speed of light. A particularly stringent test of this propagation speed is made possible by the near-coincident detection

by the Fermi gamma-ray telescope $1.74 \pm .05$ seconds after the gravitational wave signal [131]. Using the low end of the measured luminosity distance to the neutron star binary and the difference in the time of arrival of the gravitational wave and the gamma-ray burst, the constraint on the ratio of the speed of gravity to the speed of light was found to be [131],

$$-3 \cdot 10^{-15} \leq \frac{\Delta v}{v_{\rm EM}} \leq +7 \cdot 10^{-16}. \tag{2.26}$$

This limit harshly constraints alternative theories of gravity which predict significant difference in the speed of gravitational propagation [132, 133].

The status of gravitational wave observations as standard sirens permits also a direct measurement of the redshift as a function of luminosity distance. This is performed by comparing the data from the gravitational waveform to measurements of the the shift of the emission lines of the elements from the neutron star ejecta. These measurements give an entirely new measurement of the Hubble constant, independent of any previous astronomical measurements of distance. The result from the analysis of the electromagnetic counterparts found a value of the Hubble constant of $70^{+12.0}_{-8.0}$ km s^{-1} Mpc^{-1} [131]. The result is consistent with existing measurements from Planck and SHoES [134, 135], although the bounds still broad as compared those from CMB and supernovae observations.

2.2.2.4 Future tests from more numerous events

Bringing Advanced LIGO to design sensitivity will also significantly improve both the rate of detecting events similar to those discussed above and the chance of detecting novel effects. Primarily, all of the above tests used for the single neutron star event act as exciting proofs-of-concepts of what can be done with this type of data. Each of the results will continue to sharpen with additional events. Particularly enticing is the possibility that a large collection of similar multi-messenger events will offer a sharp, independent measurement of the Hubble constant, which would be particularly valuable in light of the mild tension between the results published by Planck [134] and SHoES [135] electromagnetic observations.

Further observations also have the possibility of providing qualitatively distinct effects, such as the gravitational wave signal from a long-lived supra or hyper-massive neutron star remnant. There

is also the opportunity to take advantage of the long-duration waveforms from more binary neutron stars mergers to better constrain the Post-Newtonian parameters for alternative theories of gravity.

The measured Neutron star gravitational wave signal was consistent with non-spinning neutron stars. With a large ensemble of neutron star observations in the catalog, conclusions could start to be drawn about the formation channels for neutron stars in compact binaries. Such information about the neutron star spins could then provide information about the rate of the magnetic field decay, which is relevant for the dynamics of neutron star interiors [136].

2.3 Predictions of EMRI detection in space-based detectors

The black hole binaries visible to ground based gravitational-wave detectors will be found primarily in mass range $1-100\,M_\odot$. Late in the inspiral, binary compact objects radiate gravitational waves with characteristic wavelengths comparable to the characteristic scale of the total mass of the system, m_1+m_2 . At high mass ratios, this is approximately the mass of the heavier companion. The high-sensitivity region of a gravitational wave interferometer is largely determined by the arm length of the detector, so ground based detectors cannot be constructed to be sensitive to the wavelengths of supermassive black holes, $M\sim 10^5-10^9\ M_\odot$, which can reach millions of kilometers. Instead, space-based detectors such as LISA (described broadly in Section 2.1.8.1) will be required to detect strong field effects of binaries involving supermassive black holes.

I focus primarily on the extreme mass ratio inspiral (EMRI) events that occur when a black hole with $m_2\sim 1-10\ M_\odot$ inspirals to a supermassive black hole. These events are particularly difficult to model using numerical relativity, and instead demand self-force computation, which I discuss in Section 1.3.

In this introduction section, I discuss the case for EMRIs as important and exciting events for gravitational physics (Section 2.3.1). Next, I discuss the best current anticipation of the rate of these events, which involves the known supermassive black holes and estimates of the population density of possible lighter companions (Section 2.3.2. Finally, I describe the ongoing efforts to address the challenges in extracting these events from the data measured by space-based detectors (Section 2.3.3).

2.3.1 Probing high-mass black holes with EMRIs

The predictions of general relativity are inflexible with regard to the degrees of freedom permitted in black hole spacetimes. The prediction of the few types of stable black hole spacetimes is known as the "no hair theorem" [137], and states that black holes in general relativity are permitted only three degrees of freedom: mass, spin, and electromagnetic charge. The charge of a black hole is typically disregarded by noting that current understanding of astrophysical progenitors for black holes are charge-neutral to very high precision. Therefore, astrophysical black holes should have only mass and spin.

The inspiral of stellar mass compact objects into supermassive black holes permits an important test of this prediction of general relativity. The extreme mass ratio ensures that even small deviations from the "no hair" scenario would significantly affect the dynamics of the small secondary [138, 139]. Further, as the extreme mass ratio inspiral orbits will not typically circularize [140], the full intricacies of the triperiodic Kerr orbit, interacting with high multipoles of the large companion's metric would be encoded in the emitted gravitational wave signals. The current estimates predict that the masses and spins of the supermassive black holes could be measured to $\sim 10^{-4}$ fractional uncertainty.

Initial tests of the ability to extract the detailed map of the supermassive black hole spacetime suggest that EMRIs will provide an unprecedented method to test the strong-field dynamics of general relativity [138]. Extreme mass ratio inspirals are a particularly sensitive system, which lingers in the strong-field region for a number of orbits comparable to the mass ratio. The long-duration signal and fine sensitivity to the system dynamics will provide a valuable method to test general relativity and its alternatives [141]. Full utility of the data from space-based detectors will require detailed predictions, both from general relativity and alternative theories of gravity. The development of general relativistic predictions are well underway, and discussed in detail in section (1.3), but waveforms that would be generated by inspirals in alternative theories have seen comparatively little development thus far [142].

2.3.2 Population estimates and event rate

It is largely accepted that there exist a population of supermassive black holes, each near the centers of elliptical and spiral galaxies [143–145]. The collection of supermassive black holes provides robust justification for the anticipation of a high event rate of extreme mass ratio events within the LISA detection volume. More precise estimates of the supermassive black hole population have been constructed from simulations of cosmic evolution and mergers [146, 147]. An interesting prediction from the synthesis models is that the mass accretion necessary to construct the supermassive black hole tends to create near-extremal supermassive black holes, which could be thoroughly tested by the observation of a population of extreme mass ratio inspirals.

Using the additional dynamics of density of compact objects in the galactic 'cusp' following the merger of massive black holes, estimates have been developed for the rate of extreme mass ratio inspirals [148]. The uncertainty in the models used or the prediction gives rise to a wide range in the anticipated event rate, which ranges from 1 to 2000 per year. The conservative estimates arise primarily from models which assume either a particularly pessimistic distribution of supermassive black holes from a phenomenological model [149], or a very large rate of direct plunges of small compact objects into the massive black hole, leaving few to inspiral slowly [148]. The optimistic subset of models considered in [148], in which a self-consistent synthesis model of supermassive black holes is used, and the number of plunges per inspiral is 10 or less, the prediction for the detected event rate is at least 100 per year.

2.3.3 Anticipated signal strength and data analysis

The LISA space based gravitational wave detector will be sensitive to a wide variety of sources, several of which are anticipated to be plentiful, including compact white dwarf binaries and compact objects visible by LIGO in very early parts of their inspirals, as well as binaries involving massive black holes [74]. This multitude of sources suggests that a significant part of the task of LISA data analysis will be the resolution of individual sources from the overlapping tumult of superposed gravitational wave signals.

One of the principal methods for extracting the EMRI signals from the simultaneous gravita-

tional wave signals in the same band is the same matched filtering used for LIGO data analysis [74], which is discussed in the context of a Wiener optimal filter in section 2.1.3. A valuable community effort known as the mock LISA data challenges [150,151] has made significant progress in developing the tools necessary to make best use of the LISA data. The results of these challenges indicate that EMRI waveforms with signal-to-noise ratio in excess of 20 can reliably be extracted from LISA data [151].

Chapter 2 Bibliography

[55] W. H. Press and K. S. Thorne, *Gravitational-wave astronomy*, Ann. Rev. Astron. Astrophys. **10** (1972), pp. 335–374

[56] A. Abramovici *et al.*, *LIGO: The Laser interferometer gravitational wave observatory*, Science **256** (1992), pp. 325–333

[57] C. J. Moore, R. H. Cole and C. P. L. Berry, *Gravitational-wave sensitivity curves*, Class. Quant. Grav. **32** (2015) (1), p. 015014

[58] S. T. McWilliams, J. P. Ostriker and F. Pretorius, *The imminent detection of gravitational waves from massive black-hole binaries with pulsar timing arrays* (2012)

[59] J. Aasi *et al.*, *Advanced LIGO*, Class.Quant.Grav. **32** (2015), p. 074001

[60] K. Thorne and R. Blandford, *Applications of Classical Physics* (2013)

[61] LIGO Scientific Collaboration and Virgo Collaboration, *Status of the Advanced LIGO and Advanced Virgo detectors* (2017)

[62] LIGO Scientific Collaboration and Virgo Collaboration, *GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence*, Astrophys. J. **851** (2017) (2), p. L35

[63] K. S. Thorne, *Gravitational waves* (1995)

[64] V. Braginsky, M. Gorodetsky and S. Vyatchanin, *Thermodynamical fluctuations and photothermal shot noise in gravitational wave antennae*, Phys.Lett. **A264** (1999), p. 1

[65] Y. T. Liu and K. S. Thorne, *Thermoelastic noise and homogeneous thermal noise in finite sized gravitational wave test masses*, Phys.Rev. **D62** (2000), p. 122002

[66] M. Evans, S. Ballmer, M. Fejer, P. Fritschel, G. Harry *et al.*, *Thermo-optic noise in coated mirrors for high-precision optical measurements*, Phys.Rev. **D78** (2008), p. 102003

[67] G. M. Harry, M. R. Abernathy, A. E. Becerra-Toledo, H. Armandula, E. Black *et al.*, *Titania-doped tantala/silica coatings for gravitational-wave detection*, Class.Quant.Grav. **24** (2007), pp. 405–416

[68] T. T. Lyons, M. W. Regehr and F. J. Raab, *Shot noise in gravitational-wave detectors with Fabry–Perot arms*, Appl. Opt. **39** (2000) (36), pp. 6761–6770,
URL: http://ao.osa.org/abstract.cfm?URI=ao-39-36-6761

[69] K. Somiya, *Detector configuration of KAGRA: The Japanese cryogenic gravitational-wave detector*, Class. Quant. Grav. **29** (2012), p. 124007

[70] D. Blair *et al.*, *Gravitational wave astronomy: the current status*, Sci. China Phys. Mech. Astron. **58** (2015), p. 120402

[71] Y. Michimura, K. Komori, A. Nishizawa, H. Takeda, K. Nagano, Y. Enomoto, K. Hayama, K. Somiya and M. Ando, *Particle swarm optimization of the sensitivity of cryogenic gravitational wave detector* (2018)

[72] B. Iyer, T. Souradeep, C. Unnikrishnan, S. Dhurandhar, S. Raja and A. Sengupta, *LIGO-India, Proposal of the Consortium for Indian Initiative in Gravitational-Wave Observations* (2011)

[73] S. Fairhurst, *Improved source localization with LIGO India*, J. Phys. Conf. Ser. **484** (2014), p. 012007

[74] G. O. A. Team, *The ESA-L3 Gravitational Wave Mission* (2016)

[75] J. C. Livas, *Status of Space-based Gravitational-wave Observatories (SGOs)*, Phys. Procedia **61** (2015), pp. 648–653

[76] H. Audley *et al.*, *Laser Interferometer Space Antenna* (2017)

[77] C. Trenkel, D. Wealthy, N. Dunbar, C. Warren, A. Schleicher, T. Ziegler, N. Brandt and R. Gerndt, *The Engineering of LISA Pathfinder âĂŞ the quietest Laboratory ever flown in Space*, J. Phys. Conf. Ser. **840** (2017) (1), p. 012006

[78] M. Armano *et al.*, *Beyond the Required LISA Free-Fall Performance: New LISA Pathfinder Results down to 20âĂĽâĂĽµHz*, Phys. Rev. Lett. **120** (2018) (6), p. 061101

[79] N. Seto, S. Kawamura and T. Nakamura, *Possibility of direct measurement of the acceleration of the universe using 0.1-Hz band laser interferometer gravitational wave antenna in space*, Phys. Rev. Lett. **87** (2001), p. 221103

[80] S. Kawamura *et al.*, *The Japanese space gravitational wave antenna: DECIGO*, Class. Quant. Grav. **28** (2011), p. 094011

[81] S. Sato *et al.*, *The status of DECIGO*, J. Phys. Conf. Ser. **840** (2017) (1), p. 012010

[82] LIGO Scientific Collaboration, *Binary Black Hole Mergers in the first Advanced LIGO Observing Run*, Phys. Rev. **X6** (2016) (4), p. 041015

[83] LIGO Scientific Collaboration and Virgo Collaboration, *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, Phys. Rev. Lett. **119** (2017) (16), p. 161101

[84] LIGO Scientific Collaboration and Virgo Collaboration, *GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2*, Phys. Rev. Lett. **118** (2017) (22), p. 221101

[85] LIGO Scientific Collaboration and Virgo Collaboration, *GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence*, Phys. Rev. Lett. **119** (2017) (14), p. 141101

[86] B. F. Schutz, *Gravitational Wave Astronomy: Delivering on the Promises* (2018), [Phil. Trans. Roy. Soc. Lond.A376,20170279(2018)]

[87] LIGO Scientific Collaboration, *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116** (2016) (6), p. 061102

[88] C. Cutler and K. S. Thorne, *An Overview of gravitational wave sources*, in *Proceedings, 16th International Conference on General Relativity and Gravitation: Durban, South Africa, July 15-21, 2001* (2013), pp. 72–111

[89] P. B. Graff, A. Buonanno and B. S. Sathyaprakash, *Missing Link: Bayesian detection and measurement of intermediate-mass black-hole binaries*, Phys. Rev. D **92** (2015), p. 022002, URL: https://link.aps.org/doi/10.1103/PhysRevD.92.022002

[90] P. Kumar, T. Chu, H. Fong, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, M. A. Scheel and B. Szilagyi, *Accuracy of binary black hole waveform models for aligned-spin binaries*, Phys. Rev. **D93** (2016) (10), p. 104050

[91] C. L. Fryer, *Mass limits for black hole formation*, Astrophys. J. **522** (1999), p. 413

[92] H. Aihara *et al.*, *The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III*, Astrophys. J. Suppl. **193** (2011), p. 29, [Erratum: Astrophys. J. Suppl.195,26(2011)]

[93] J. C. Niemeyer and K. Jedamzik, *Dynamics of primordial black hole formation*, Phys. Rev. **D59** (1999), p. 124013

[94] K. Belczynski, V. Kalogera and T. Bulik, *A Comprehensive study of binary compact objects as gravitational wave sources: Evolutionary channels, rates, and physical properties*, Astrophys. J. **572** (2001), pp. 407–431

[95] M. Dominik, E. Berti, R. O'Shaughnessy, I. Mandel, K. Belczynski, C. Fryer, D. E. Holz, T. Bulik and F. Pannarale, *Double Compact Objects III: Gravitational Wave Detection Rates*, Astrophys. J. **806** (2015) (2), p. 263

[96] C. L. Rodriguez, M. Morscher, B. Pattabiraman, S. Chatterjee, C.-J. Haster and F. A. Rasio, *Binary Black Hole Mergers from Globular Clusters: Implications for Advanced LIGO*, Phys. Rev. Lett. **115** (2015) (5), p. 051101, [Erratum: Phys. Rev. Lett.116,no.2,029901(2016)]

[97] J. Abadie *et al.*, *Search for Gravitational Waves from Intermediate Mass Binary Black Holes*, Phys. Rev. **D85** (2012), p. 102004

[98] V. Cardoso, L. Gualtieri, C. Herdeiro and U. Sperhake, *Exploring New Physics Frontiers Through Numerical Relativity*, Living Rev. Relativity **18** (2015), p. 1

[99] T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, *Modified Gravity and Cosmology*, Phys. Rept. **513** (2012), pp. 1–189

[100] J.-P. Bruneton and G. Esposito-Farese, *Field-theoretical formulations of MOND-like gravity*, Phys. Rev. **D76** (2007), p. 124012, [Erratum: Phys. Rev.D76,129902(2007)]

[101] T. P. Sotiriou and V. Faraoni, *f(R) Theories Of Gravity*, Rev. Mod. Phys. **82** (2010), pp. 451–497

[102] M. Okounkova, L. C. Stein, M. A. Scheel and D. A. Hemberger, *Numerical binary black hole mergers in dynamical Chern-Simons gravity: Scalar field*, Phys. Rev. **D96** (2017) (4), p. 044020

[103] LIGO Scientific Collaboration and Virgo Collaboration, *Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA* (2013), [Living Rev. Rel.19,1(2016)]

[104] B. Carr, M. Raidal, T. Tenkanen, V. Vaskonen and H. VeermÃďe, *Primordial black hole constraints for extended mass functions*, Phys. Rev. **D96** (2017) (2), p. 023514

[105] J. G. Rosa, *Boosted black string bombs*, JHEP **02** (2013), p. 014

[106] W. H. Press and S. A. Teukolsky, *Floating Orbits, Superradiant Scattering and the Black-hole Bomb*, Nature **238** (1972), pp. 211–212

[107] J. Arons, *Magnetars in the metagalaxy: an origin for ultrahigh-energy cosmic rays in the nearby universe*, Astrophys. J. **589** (2003), pp. 871–892

[108] S. A. Olausen and V. M. Kaspi, *The McGill Magnetar Catalog*, Astrophys. J. Suppl. **212** (2014), p. 6

[109] M. Soares-Santos *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/Virgo GW170817. I. Discovery of the Optical Counterpart Using the Dark Energy Camera*, Astrophys. J. **848** (2017) (2), p. L16

[110] P. S. Cowperthwaite *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/Virgo GW170817. II. UV, Optical, and Near-infrared Light Curves and Comparison to Kilonova Models*, Astrophys. J. **848** (2017) (2), p. L17

[111] M. Nicholl *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. III. Optical and UV Spectra of a Blue Kilonova From Fast Polar Ejecta*, Astrophys. J. **848** (2017) (2), p. L18

[112] R. Chornock *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. IV. Detection of Near-infrared Signatures of r-process Nucleosynthesis with Gemini-South*, Astrophys. J. **848** (2017) (2), p. L19

[113] R. Margutti *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. V. Rising X-ray Emission from an Off-Axis Jet*, Astrophys. J. **848** (2017) (2), p. L20

[114] K. D. Alexander *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. VI. Radio Constraints on a Relativistic Jet and Predictions for Late-Time Emission from the Kilonova Ejecta*, Astrophys. J. **848** (2017) (2), p. L21

[115] P. K. Blanchard *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. VII. Properties of the Host Galaxy and Constraints on the Merger Timescale*, Astrophys. J. **848** (2017) (2), p. L22

[116] W. Fong *et al.*, *The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. VIII. A Comparison to Cosmological Short-duration Gamma-ray Bursts*, Astrophys. J. **848** (2017) (2), p. L23

[117] K. Rajagopal and F. Wilczek, *The Condensed matter physics of QCD*, in M. Shifman and B. Ioffe (eds.), *At the frontier of particle physics. Handbook of QCD. Vol. 1-3* (2000), pp. 2061–2151

[118] K. Schertler, C. Greiner, J. Schaffner-Bielich and M. H. Thoma, *Quark phases in neutron stars and a 'third family' of compact stars as a signature for phase transitions*, Nucl. Phys. **A677** (2000), pp. 463–490

[119] F. ÃŰzel and P. Freire, *Masses, Radii, and the Equation of State of Neutron Stars*, Ann. Rev. Astron. Astrophys. **54** (2016), pp. 401–440

[120] LIGO Scientific Collaboration and Virgo Collaboration, *Multi-messenger Observations of a Binary Neutron Star Merger*, Astrophys. J. **848** (2017) (2), p. L12

[121] T. Hinderer, B. D. Lackey, R. N. Lang and J. S. Read, *Tidal deformability of neutron stars with realistic equations of state and their gravitational wave signatures in binary inspiral*, Phys. Rev. **D81** (2010), p. 123016

[122] T. Binnington and E. Poisson, *Relativistic theory of tidal Love numbers*, Phys. Rev. **D80** (2009), p. 084018

[123] S. E. Gralla, *On the Ambiguity in Relativistic Tidal Deformability*, Class. Quant. Grav. **35** (2018) (8), p. 085002

[124] S. Goriely, *Radiative neutron captures by neutron-rich nuclei and the r-process nucleosynthesis*, Phys. Lett. **B436** (1998), pp. 10–18

[125] F. K. Thielemann, T. Rauscher, C. Freiburghaus, K. Nomoto, M. Hashimoto, B. Pfeiffer and K. L. Kratz, *Nucleosynthesis basics and applications to supernovae* (1998)

[126] O. Korobkin, S. Rosswog, A. Arcones and C. Winteler, *On the astrophysical robustness of neutron star merger r-process*, Mon. Not. Roy. Astron. Soc. **426** (2012), p. 1940

[127] F. K. Thielemann, D. Argast, F. Brachwitz, J. L. Fisker, C. FrÃűhlich, R. Hirschi, E. Kolbe, D. Mocelj and T. Rauscher, *Nuclear Physics: A Key Ingredient in Astrophysical Modeling*, Nucl. Phys. **A751** (2005), pp. 301–326

[128] M. Arnould, S. Goriely and K. Takahashi, *The r-process of stellar nucleosynthesis: Astrophysics and nuclear physics achievements and mysteries*, Phys. Rept. **450** (2007), pp. 97–213

[129] K. Takami, L. Rezzolla and L. Baiotti, *Spectral properties of the post-merger gravitational-wave signal from binary neutron stars*, Phys. Rev. **D91** (2015) (6), p. 064001

[130] LIGO Scientific Collaboration and Virgo Collaboration, *Search for Post-merger Gravitational Waves from the Remnant of the Binary Neutron Star Merger GW170817*, Astrophys. J. **851** (2017) (1), p. L16

[131] LIGO Scientific Collaboration and Virgo Collaboration, *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, Astrophys. J. **848** (2017) (2), p. L13

[132] J. Sakstein and B. Jain, *Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories*, Phys. Rev. Lett. **119** (2017) (25), p. 251303

[133] S. Nojiri and S. D. Odintsov, *Cosmological Bound from the Neutron Star Merger GW170817 in scalar-tensor and $F(R)$ gravity theories*, Phys. Lett. **B779** (2018), pp. 425–429

[134] N. Aghanim *et al.*, *Planck intermediate results. XLVI. Reduction of large-scale systematic effects in HFI polarization maps and estimation of the reionization optical depth*, Astron. Astrophys. **596** (2016), p. A107

[135] A. G. Riess *et al.*, *A 2.4% Determination of the Local Value of the Hubble Constant*, Astrophys. J. **826** (2016) (1), p. 56

[136] X. Zhu, E. Thrane, S. Oss, Y. Levin and P. D. Lasky, *Inferring the population properties of binary neutron stars with gravitational-wave measurements of spin* (2017)

[137] E. Berti, V. Cardoso and C. M. Will, *On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA*, Phys. Rev. **D73** (2006), p. 064030

[138] L. Barack and C. Cutler, *Using LISA EMRI sources to test off-Kerr deviations in the geometry of massive black holes*, Phys. Rev. **D75** (2007), p. 042003

[139] S. A. Hughes, *Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational wave emission. II. Inspiral trajectories and gravitational wave forms*, Phys. Rev. **D64** (2001), p. 064004, [Erratum: Phys. Rev.D88,no.10,109902(2013)]

[140] S. Drasco and S. A. Hughes, *Gravitational wave snapshots of generic extreme mass ratio inspirals*, Phys. Rev. **D73** (2006) (2), p. 024027, [Erratum: Phys. Rev.D90,no.10,109905(2014)]

[141] J. R. Gair, M. Vallisneri, S. L. Larson and J. G. Baker, *Testing General Relativity with Low-Frequency, Space-Based Gravitational-Wave Detectors*, Living Rev. Rel. **16** (2013), p. 7

[142] P. Canizares, J. R. Gair and C. F. Sopuerta, *Testing Chern-Simons modified gravity with observations of extreme-mass-ratio binaries*, J. Phys. Conf. Ser. **363** (2012), p. 012019

[143] K. Gultekin *et al.*, *The M-sigma and M-L Relations in Galactic Bulges and Determinations of their Intrinsic Scatter*, Astrophys. J. **698** (2009), pp. 198–221

[144] R. Genzel, F. Eisenhauer and S. Gillessen, *The Galactic Center Massive Black Hole and Nuclear Star Cluster*, Rev. Mod. Phys. **82** (2010), pp. 3121–3195

[145] J. Kormendy and K. Gebhardt, *Supermassive black holes in nuclei of galaxies*, AIP Conf. Proc. **586** (2001), pp. 363–381

[146] E. Barausse, *The evolution of massive black holes and their spins in their galactic hosts*, Mon. Not. Roy. Astron. Soc. **423** (2012), pp. 2533–2557

[147] F. Antonini, E. Barausse and J. Silk, *The Coevolution of Nuclear Star Clusters, Massive Black Holes, and their Host Galaxies*, Astrophys. J. **812** (2015) (1), p. 72

[148] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. L. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau and A. Klein, *Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals*, Phys. Rev. **D95** (2017) (10), p. 103012

[149] J. R. Gair, C. Tang and M. Volonteri, *LISA extreme-mass-ratio inspiral events as probes of the black hole mass function*, Phys. Rev. **D81** (2010), p. 104014

[150] S. Babak *et al.*, *The Mock LISA Data Challenges: From Challenge 1B to Challenge 3*, Class. Quant. Grav. **25** (2008), p. 184026

[151] S. Babak *et al.*, *The Mock LISA Data Challenges: From Challenge 3 to Challenge 4*, Class. Quant. Grav. **27** (2010), p. 084009

Chapter 3

Radiation Reaction of Small Charged Bodies to Second Order

Coauthor:

Éanna Flanagan, Cornell University

Published: Phys. Rev. D 97, 105001

3.1 Introduction

3.1.1 Status of our understanding of self force effects

Classical electrodynamics dictates that an accelerating charge emits radiation. This electromagnetic radiation carries energy and momentum, so conservation laws demand that the charge must experience a force. The force arises from the charge interacting with its own field, and is known as the 'radiation-reaction force' or 'self force'. This phenomenon was first derived by Lorentz [152], and later confirmed by Abraham [153] followed by Dirac [154], each expanding and generalizing the results of the prior work.

Computing expressions for self forces is notoriously complicated, and there is an enormous literature on this field. The complexity arises in part because self forces describe back-reaction: as a charge accelerates, its radiation perturbs its motion, in turn altering the details of the radiation. Analytic methods are tractable in the regime in which the body is small compared to the characteristic lengthscales of the external fields. In this limit, the self force can be expanded order by order in the charge of the body. In this paper, we use the common nomenclature of referring to the Lorentz

force as the leading order force, the leading correction to the Lorentz force as the 'first order' self force, and so on. Our understanding of radiation reaction in flat spacetime has been developed over most of a century [155–158], culminating in the rigorous treatment of Gralla, Harte, and Wald [159] (henceforth GHW) who carefully analyzed a limit in which the charge, size, and mass of a body go to zero. The modern focus of the self force community is that of small masses in curved spacetime, for which Eric Possion's review article offers a thorough introduction [160].

The self force is of great interest to modern astrophysics. Just as a charged particle interacts with its own field as it radiates electromagnetic waves, gravitating systems experience self forces from the emission of gravitational radiation. The gravitational waves produced by binary black hole inspirals and binary neutron star inspirals have been detected by LIGO [161, 162], and similar binary inspirals are candidate signals for the future space-based detector LISA.

Making full use of the data from LISA will require an improved understanding of self force effects. The gravitational self force to leading order in the mass of the small body is referred to as the MiSaTaQuWa self force, and was first derived in [163, 164]. More recent computations have extended these results to second order [165–170], and applied the self force to a gravitational inspiral, in order to compute or numerically evaluate the worldline and the resulting gravitational radiation. The computational strategies for evaluating worldlines and waveforms from gravitational self force are reviewed well in [171, 172]. The techniques for computing leading order, or adiabatic, waveforms are now known. However, LISA data analysis will require post-adiabatic waveform predictions, which in turn will also require the subleading self force. This motivates a detailed understanding of the subleading self force.

Previous derivations of higher-order self forces for non-gravitational fields include those of Chad Galley [173] and Abraham Harte [174]. Galley's derivation [173] of the scalar self force uses an effective field technique to derive the self force to high order for monopolar charges. Harte has derived exact expressions for the self force of an extended charge distribution in an external field. The relation between Harte's results and our work is somewhat involved and is discussed in Sec. 3.3 below.

3.1.2 The Gralla-Harte-Wald derivation method and its extension

In this paper, we derive the subleading order electromagnetic and scalar self forces acting on a small charged body moving in flat spacetime. The calculation is motivated by the importance of the gravitational self force, and is a model for the more complicated computation in the gravitational case. Although subleading self forces have previously been computed [175, 176], ours is the first to describe extended body effects to subleading order. In addition to providing a model for the gravitational self force, our calculation may have direct application to systems with extremely strong electromagnetic fields, as discussed further below.

GHW introduce a one-parameter family of bodies with the property that as the parameter approaches zero, the mass, charge, and spatial extent of the body approach zero at the same rate. By considering various moments of the stress-energy conservation and charge conservation equations, integrated over a small region containing the body, they derive the first-order self force, mass evolution, and spin evolution equations.

Our calculation uses the GHW axioms with slight modifications, which are presented in full in section 3.4. However, we found it necessary to modify and refine the definitions of body parameters. GHW defined parameters such as the total mass-energy, angular momentum, and electromagnetic multipole moments in terms of integrals over a spacelike hypersurface perpendicular to the center of mass worldline[1]. At second order, these definitions are problematic, and we replace them with body parameter definitions in terms of integrals over the future null cones of points on the center of mass worldline. With these definitions, the body parameters at a given time depend only on the body's stress-energy and charge distribution at times within a light crossing time, not on the stress-energy or charge distribution in the distant past. This is because, in flat spacetime, the field at every point depends only on sources on that point's past lightcone.

[1] As usual, there are ambiguities in the precise definition of center of mass worldline [174]. These ambiguities affect the form of the equation of motion at subleading orders, and are associated with the choice of a spin supplementary condition. See Section 3.2.2 below.

3.1.3 Discussion of results - applications in physical systems

Our results for the second order evolution of the body's worldline, mass, and spin are given in Eqs. (3.75) - (3.78). They contain three types of terms: coupling of electromagnetic moments to the external field, self force terms that do not depend on the higher electromagnetic moments, and terms which describe a mixing between self-field and extended body effects. Our spin evolution equation contains a self-torque, which was not seen previously at lower orders. Our results also satisfy a consistency check obtained by comparing with some non-perturbative results of Harte [174].

As an illustrative special case, consider a body with vanishing spin, electromagnetic dipole, and quadrupole, moving in an external electromagnetic field $F^{(\mathrm{ext})\mu\nu}$. The acceleration of the body can be written as [c.f. Eq. (3.81) below], in units with $c = 1$,

$$\begin{aligned} a^\mu =& \kappa F^{(\mathrm{ext})\mu\lambda} u_\lambda + q\left\{\frac{2}{3}\kappa^2 D_\tau F^{(\mathrm{ext})\mu\lambda} u_\lambda + \frac{2}{3}\kappa^3 \mathcal{P}^\mu{}_\nu F^{(\mathrm{ext})\nu\lambda} F^{(\mathrm{ext})}{}_{\lambda\sigma} u^\sigma\right\} \\ &+ q^2\left\{\frac{4}{9}\kappa^3 D_\tau{}^2 F^{(\mathrm{ext})\mu\lambda} u_\lambda + \frac{8}{9}\kappa^4 \mathcal{P}^\mu{}_\kappa F^{(\mathrm{ext})\kappa\lambda} F^{(\mathrm{ext})}{}_{\lambda\sigma} u^\sigma \right. \\ &\left. + \frac{4}{9}\kappa^4 \mathcal{P}^\mu{}_\kappa F^{(\mathrm{ext})\kappa\lambda} D_\tau F^{(\mathrm{ext})}_{\lambda\sigma} u^\sigma + \frac{4}{9}\kappa^5 \mathcal{P}^\mu{}_\kappa F^{(\mathrm{ext})\kappa\rho} \mathcal{P}_{\rho\lambda} F^{(\mathrm{ext})\lambda\sigma} F^{(\mathrm{ext})}{}_{\sigma\omega} u^\omega\right\} + \mathcal{O}(q^3). \end{aligned} \tag{3.1}$$

Here u^μ is the 4-velocity of the body, a^μ the 4-acceleration, $D_\tau \equiv u^\mu \nabla_\mu$, and $\mathcal{P}^\mu{}_\nu = \delta^\mu{}_\nu + u^\mu u_\nu$ is the projection tensor. Also, q is the charge and $\kappa = q/m$ is the charge to mass ratio. The right hand side consists of an expansion in q at fixed κ. The first term is the Lorentz force law, the second term is the reduced-order (see Sec. 3.5.1 below) form of the Abraham-Lorentz-Dirac equation, and the third term is our new result.

We now turn to a discussion of the domain of validity of our results. Consider a charged body of mass m, and charge q, moving in an external field that imparts a characteristic acceleration a, as measured in the body's instantaneous rest-frame. Suppose also that the field varies on some timescale or lengthscale τ_{ext}, again as measured in the body's instantaneous rest-frame. Then there are a number of conditions that must be satisfied for our analysis to be valid:

- *Small multipole couplings*: If the condition

$$\mathcal{R} \ll \tau_{\mathrm{ext}}, \tag{3.2}$$

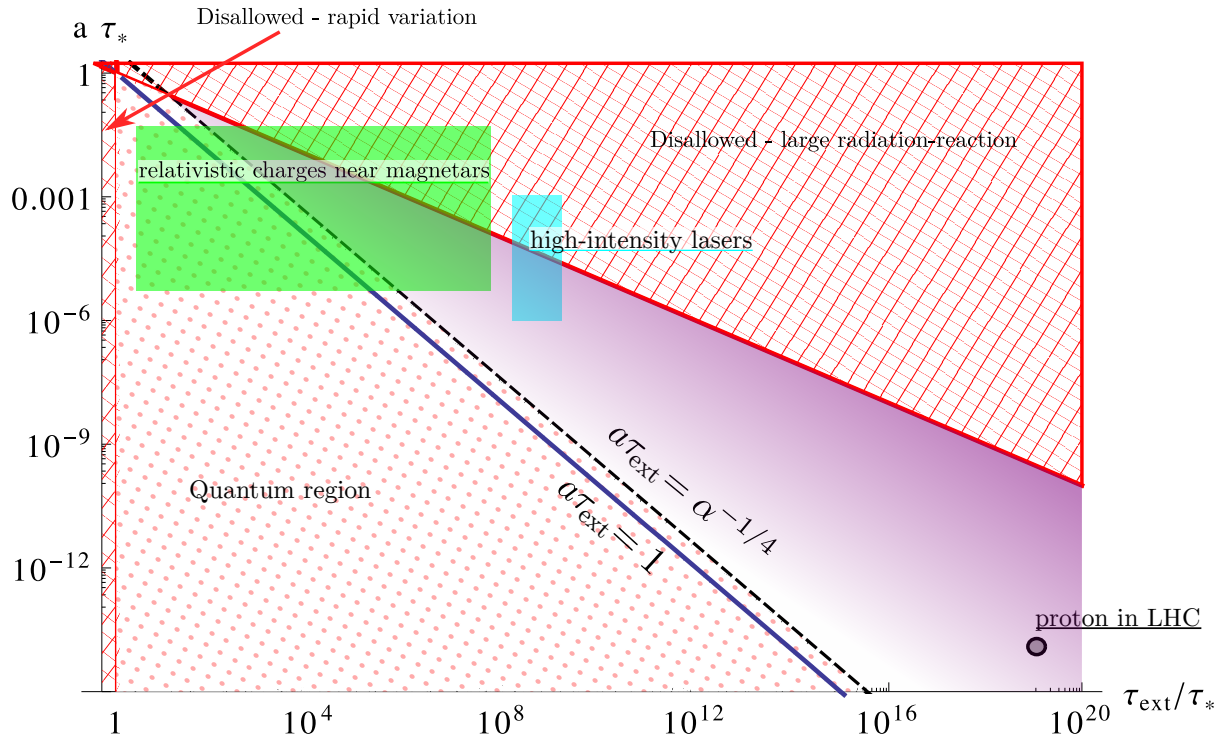

Figure 3.1: An illustration of the parameter space for radiation reaction for relativistic particles. The horizontal axis is the ratio $\tau_{\rm ext}/\tau_*$, where $\tau_{\rm ext}$ is the timescale over which the external field is varying, as measured in the instantaneous rest frame of the particle, and $\tau_* = q^2/m$, where q is the charge and m the mass of the particle. The vertical axis is $a\tau_*$, where a is the acceleration due to the external field. The motion is relativistic in the region $a\tau_{\rm ext} \gg 1$, in the upper right hand of the figure. Radiation reaction effects are large in the shaded regions where $\tau_*/\tau_{\rm ext} \gtrsim 1$ or $a^2\tau_{\rm ext}\tau_* \gtrsim 1$. These regions lie outside of the domain of validity of our analysis, and the second order self force is negligible except near the boundaries of these regions. In the region below and to the left of the dashed line, the radiation from the particle is not in a classical regime and our analysis does not apply. We show the location of protons in the Large Hadron Collider, protons in very high intensity lasers, and electrons in the high magnetic fields of magnetars.

is satisfied then the leading order couplings (dipole, quadrupole, and so on) will dominate.

- *Weak radiation reaction*: The energy radiated in a dynamical time must be small compared to the change in the body's energy due to conservative effects. If this is violated then our derivation is no longer valid. In the non-relativistic region $a\tau_{\text{ext}} \ll 1$ this requires

$$\frac{\tau_*}{\tau_{\text{ext}}} \ll 1 \tag{3.3}$$

where $\tau_* = q^2/m$. In the relativistic regime $a\tau_{\text{ext}} \gg 1$, the condition is instead

$$a^2\tau_{\text{ext}}\tau_* \ll 1. \tag{3.4}$$

- *Classical radiation regime*: The energy radiated in a dynamical time must be large compared to the energy radiated per quantum, so that many quanta are emitted in a dynamical time. In the non-relativistic regime $a\tau_{\text{ext}} \ll 1$ the corresponding requirement is

$$a\tau_{\text{ext}} \gg \alpha^{-1/2}, \tag{3.5}$$

where $\alpha = q^2/\hbar$, and the relativistic regime $a\tau_{\text{ext}} \gg 1$ it is

$$a\tau_{\text{ext}} \gg \alpha^{-1/4}. \tag{3.6}$$

For elementary particles typically $\alpha \ll 1$ while for macroscopic charged bodies $\alpha \gg 1$.

Our derivation method employs a certain limiting procedure which automatically enforces the conditions (3.2),(3.3), and (3.4). The two dimensional parameter space of acceleration a and external timescale τ_{ext} is illustrated in Fig 3.1. The solid line $a\tau_{\text{ext}} = 1$ is the boundary between non-relativistic and relativistic motion; the lower left region is non-relativistic while the upper right is relativistic. The shaded regions on the left and at the top correspond to strong radiation reaction and lie outside our domain of validity, by (3.3) and (3.4). Our second order self force will be significant only near these boundaries. The region to the left of the dashed line is disallowed since the radiation is not classical, by (3.6) (assuming an elementary particle so that $\alpha \ll 1$). Also shown on the plot are some illustrative examples:

- A proton at the Large Hadron Collider, for which $a \sim 3 \cdot 10^{12}\,\text{s}^{-1}$, $\tau_{\text{ext}} \sim 1.4 \cdot 10^{-8}\,\text{s}$, $\tau_* \sim 6 \cdot 10^{-27}\,\text{s}$. In this case we have $a^2\tau_{\text{ext}}\tau_* \sim 10^{-9}$, so higher order radiation reaction effects are

negligible. Lead ions in the LHC experience a similar acceleration, and have a τ_* almost two orders of magnitude larger, $\tau_* \sim 2 \cdot 10^{-25}\,\mathrm{s}$, so the scale of effect is $a^2\tau_{\mathrm{ext}}\tau_* \sim 10^{-8}$.

- For high-intensity laser systems with intensities in the range $10^{19}\,\mathrm{W/cm^2} - 10^{22}\,\mathrm{W/cm^2}$ [177–179], the acceleration scale for a proton is then in the range $a \sim 10^{17}\,\mathrm{s^{-1}} - 10^{21}\,\mathrm{s^{-1}}$, and using $\tau_{\mathrm{ext}} \sim 10^{-16}\,\mathrm{s}$ and $\tau_* \sim 6 \cdot 10^{-23}\,\mathrm{s}$ gives $a^2\tau_{\mathrm{ext}}\tau_*$ in the range 10^{-8}-10^{0}. At the upper end of this range, second order radiation reaction effects could become significant. [180]

- Turning to astrophysics, the magnetic fields near certain neutron stars, referred to as "magnetars", can be extremely large, $B \sim 10^8 - 10^{11}\mathrm{T}$. At the high end of this range, higher order self force effects could easily become large even for slowly moving particles.

3.2 Motion of a finite body coupled to an external field

In this section, we consider a finite extended body moving in an external field in flat spacetime. We will review the governing equation, the non-perturbative definition of the body parameters. In the following sections we will review the non-perturbative equations of motion for the body moments, and specialize to the limit of a small body to obtain explicit results.

3.2.1 Governing equations

The system we are considering is a finite, extended, charged body coupled to an external field in flat spacetime. The extended body is described by a matter stress-energy tensor $T_M^{\mu\nu}$, which we assume is smooth and which vanishes outside a world tube of compact spatial support. We will consider both electromagnetic and scalar self forces.

The coupling to either type of field is governed by the body's charge, which is described by a charge current density j^μ such that $\nabla_\mu j^\mu = 0$ (electromagnetic case), or a scalar charge density ρ (scalar case). We assume that the charge current or density functions are also smooth and of compact spatial support. These fields obey the standard inhomogeneous wave equations for the

respective type of field:

$$\nabla_{[\mu} F_{\lambda\sigma]} = 0, \tag{3.7a}$$

$$\nabla^{\mu} F_{\mu\nu} = 4\pi j_{\nu} \qquad \text{(E\&M case)}, \tag{3.7b}$$

and

$$\nabla_{\mu}\nabla^{\mu}\Phi = -4\pi\rho \qquad \text{(scalar case)}. \tag{3.8}$$

The total stress-energy tensor $T_{\mu\nu}$ is given by the sum of the matter contribution $T_{M\,\mu\nu}$ and the field contribution $T_{F\,\mu\nu}$. This stress energy contribution for the electromagnetic field is

$$4\pi T_{F\,\mu\nu} = F_{\mu\lambda}F^{\lambda}{}_{\nu} - \frac{1}{4} g_{\mu\nu} F_{\sigma\lambda} F^{\sigma\lambda}, \tag{3.9}$$

or, for the scalar field, is

$$4\pi T_{F\,\mu\nu} = \nabla_{\mu}\Phi\nabla_{\nu}\Phi - \frac{1}{2} g_{\mu\nu}\nabla_{\lambda}\Phi\nabla^{\lambda}\Phi. \tag{3.10}$$

We assume that this total stress-energy is conserved:

$$\nabla_{\mu}\left(T^{\mu\nu}_{\mathrm{M}} + T^{\mu\nu}_{\mathrm{F}}\right) = 0. \tag{3.11}$$

We choose to divide the field into an external field $F^{(\mathrm{ext})\mu\nu}$ (Scalar: $\Phi^{(\mathrm{ext})}$), and a self field $F^{(\mathrm{self})\mu\nu}$ (Scalar: $\Phi^{(\mathrm{self})}$) which is the retarded solution to the field equations (3.7) or (3.8) with the given source. The external field may be expressed as, for the electromagnetic case,

$$F^{(\mathrm{ext})}{}_{\mu\nu} = F_{\mu\nu} - F^{(\mathrm{self})}{}_{\mu\nu}, \tag{3.12}$$

or, for the scalar case,

$$\Phi^{(\mathrm{ext})} = \Phi - \Phi^{(\mathrm{self})}. \tag{3.13}$$

Inserting the decompositions (3.12),(3.13) into the quadratic expressions (3.9),(3.10) for the field stress energy tensor, we find following GHW that the field stress energy can be expressed as the sum of three terms:

$$T^{\mu\nu}_{\mathrm{F}} = T^{\mu\nu}_{(\mathrm{self})} + T^{\mu\nu}_{(\mathrm{cross})} + T^{\mu\nu}_{(\mathrm{ext})}. \tag{3.14}$$

Here $T^{\mu\nu}_{(\mathrm{self})}$ is quadratic in the self field, $T^{\mu\nu}_{(\mathrm{ext})}$ is quadratic in the external field, and $T^{\mu\nu}_{(\mathrm{cross})}$ is a cross term which depends on both the self field and the external field.

In the following subsection we will discuss the definition of body parameters such as mass, momentum, and spin. For those definitions, we will use the sum of the matter and self stress energy tensors,

$$T^{\mu\nu} = T^{\mu\nu}_{\rm M} + T^{\mu\nu}_{\rm (self)}, \tag{3.15}$$

excluding the cross and external contribution, following GHW. The conservation of stress-energy (3.11) can be rewritten in terms of this quantity as:

$$\nabla_\mu T^{\mu\nu} = F^{{\rm (ext)}\nu\mu} j_\mu \qquad \text{(E\&M case)}, \tag{3.16a}$$

$$\nabla_\mu T^{\mu\nu} = \Phi^{{\rm (ext)};\nu} \rho \qquad \text{(scalar case)}. \tag{3.16b}$$

The motivation for choosing the definition (3.15) for the body parameter definitions is that in the limit when the body becomes small, the fields $T^{\mu\nu}$, j^μ, and ρ vary over the small body lengthscale, while the external fields $F^{{\rm (ext)}\mu\nu}$ and $\Phi^{{\rm (ext)};\mu}$ vary only on a longer lengthscale set by the external field.

The only equations that are needed for our derivation of the self force are the field equations (3.7) and (3.8), the stress energy conservation equation in the form (3.16), and the definition of the self-field as the retarded field.

3.2.2 Non-perturbative definition of body parameters: the Dixon-Harte formalism

We now turn to a discussion of the definition of body parameters for a finite body, including the body's mass, momentum, spin, and choice of representative worldline.

For a conserved stress energy tensor $T^{\mu\nu}$ in flat spacetime of compact spatial support, there is a natural choice of momentum and spin, namely

$$P^\mu_{\rm (Isolated)} = \int_\Sigma T^{\mu\nu} d\Sigma_\nu, \tag{3.17a}$$

$$S^{\mu\nu}_{\rm (Isolated)}(z^\mu) = 2\int_\Sigma (x-z)^{[\mu} T^{\nu]\lambda} d\Sigma_\lambda, \tag{3.17b}$$

where Σ is any spacelike hypersurface. The center of mass worldline is then the set of points z^μ which satisfy

$$S^{\mu\nu}_{\rm (Isolated)}(z^\mu) P_{{\rm (Isolated)}\nu} = 0. \tag{3.18}$$

Equation (3.18) is known as a spin supplementary condition, and generalizations of this condition will be discussed below.

However, this treatment is not applicable to our present context for two reasons:

- First, the stress-energy tensor (3.15) that we wish to use in the definitions is not conserved, instead there is a forcing term from the external field on the right hand side of Eqs. (3.16). Hence, the expressions (3.17) will no longer be independent of the choice of hypersurface Σ, and a specific choice of hypersurface Σ will be required. This will be discussed further below.

- Second, the stress energy term (3.15) that we will use does not have compact spatial support, due to the self field contribution. Hence, there is no guarantee that the expressions (3.17) are convergent and well defined. The convergence of these integrals is discussed further below.

There exists a general, fully non-perturbative set of definitions of worldlines, electromagnetic moments, and stress-energy moments of an extended body. These definitions were introduced by Dixon [181, 182] in the context of curved spacetime, and extended by Harte [174]. We follow the Dixon-Harte framework and definitions, with some modifications that we discuss below. The remainder of this section reviews those aspects of the Dixon-Harte framework that are most important for our derivation.

Before discussing the definitions of body parameters, we review the covariant bitensor formalism [160]. We work in flat spacetime, but we will be using non-Lorentzian coordinates. We will denote by $x^{\tilde{\mu}}$ a field point off the worldline, and we use tilded indices for tensors at such points. We will denote by $z^\mu(\tau)$ a point on the worldline (figure 3.2), and use normal (untilded) indices for the tensors at such points. General bitensors are functions of both z^μ and $x^{\tilde{\mu}}$, and can have one or more indices of either type.

An important set of bitensors are Synge's worldfunction $\sigma(x,z)$ and its derivatives. Synge's worldfunction is defined only for pairs of points that are sufficiently close that there exists a unique geodesic that joins them. For this unique geodesic, $\sigma(x,z)$ measures the half geodesic distance squared between the two points. It is negative for timelike separated points, positive for spacelike separated points, and zero for null-related points. The first covariant derivative of Synge's worldfunction can be used to define a covariant version of a position vector $\sigma_\mu(x,z) \equiv \nabla_\mu \sigma(x,z)$, where the

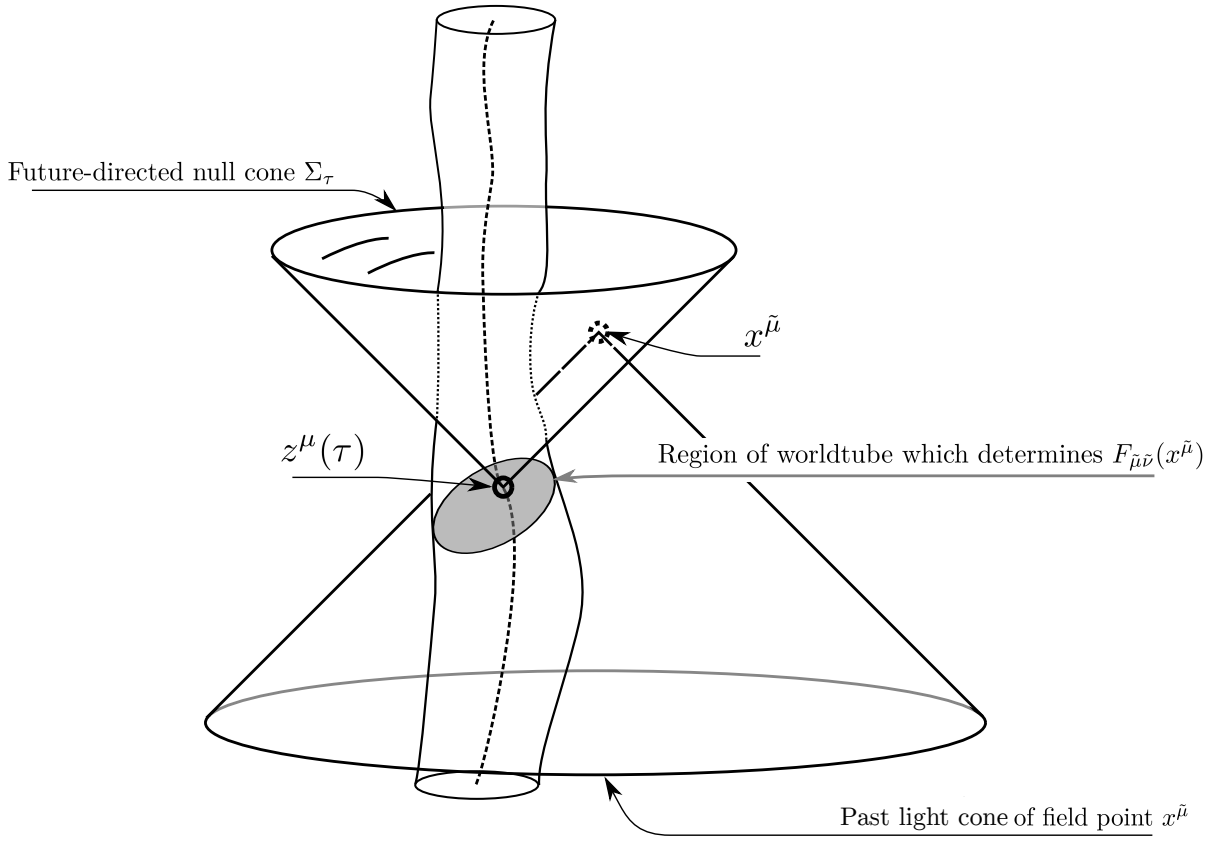

Figure 3.2: An illustration of our definitions of total momentum and spin of an extended body. The body is confined to the world tube shown, but is coupled to a long range field (scalar or electromagnetic) that extends beyond the worldtube. Given a representative worldline $z^\mu(\tau)$, shown as a dashed line, we define momentum and spin by integrating over future null cones Σ_τ of points on the worldline. The field stress energy tensor at a point $x^{\tilde{\mu}}$ on such a null cone will depend on the sources in the intersection of its past lightcone with the worldtube, shaded in gray. This region is confined to within the region of the worldtube consisting of times τ' with $|\tau - \tau'|$ smaller than a light-crossing time.

derivative is with respect to z. We will also find useful the second derivatives, $\sigma^\mu{}_\lambda(x,z) \equiv \nabla_\lambda \nabla^\mu \sigma$ and $\sigma^{\tilde{\mu}}{}_\lambda \equiv \nabla^{\tilde{\mu}} \nabla_\lambda \sigma$.

In the Dixon-Harte framework, one chooses a worldline $z^\alpha(\tau)$ for the body, where τ is a parameter that need not be proper time, and a choice of a unit vector $n^\alpha(\tau)$ along the worldline with $n_\alpha (d/d\tau)^\alpha = -1$. The formalism supplies conditions that eventually determine the worldline and parameterization. Given these choices, one defines a foliation of spacetime by hypersurfaces Σ_τ as follows. Each hypersurface is labeled by the parameter τ at which it intersects the worldline, so $z^\alpha(\tau) \in \Sigma_\tau$, and is generated by geodesics starting on the worldline that are orthogonal to n^α.

The Dixon-Harte definitions of the momentum and spin of an extended body are

$$P^{\rm D}_{\mu}(\tau) = \int_{\Sigma_\tau} d\Sigma_{\tilde{\mu}}(x) T^{\tilde{\mu}\tilde{\nu}}_{M}(x) K_{\tilde{\nu}\mu}(x, z_\tau), \tag{3.19a}$$

$$S^{\rm D}_{\mu\nu}(\tau) = 2\int_{\Sigma_\tau} d\Sigma_{\tilde{\mu}}(x) T^{\tilde{\mu}\tilde{\nu}}_{M}(x) H_{\tilde{\nu}[\mu}(x, z_\tau)\sigma_{\nu]}(x, z_\tau), \tag{3.19b}$$

where

$$H^{\tilde{\mu}}{}_{\nu} = -\left(\sigma^{\nu}{}_{\tilde{\mu}}\right)^{-1}, \tag{3.20a}$$

$$K^{\tilde{\mu}}{}_{\nu} = H^{\tilde{\mu}}{}_{\lambda}\sigma^{\lambda}{}_{\nu}. \tag{3.20b}$$

In flat spacetime, these definitions reduce to:

$$P^{\mu}_{D}(\tau) = \int_{\Sigma_\tau} d\Sigma_{\tilde{\mu}}(x^{\tilde{\lambda}}) T^{\tilde{\mu}\tilde{\nu}}_{M}(x^{\tilde{\lambda}}) g_{\tilde{\nu}}{}^{\mu}(x^{\tilde{\lambda}}, z^{\lambda}(\tau)), \tag{3.21a}$$

$$S^{\mu\nu}_{D}(\tau) = 2\int_{\Sigma_\tau} d\Sigma_{\tilde{\mu}}(x^{\tilde{\lambda}}) T^{\tilde{\mu}\tilde{\nu}}_{M}(x^{\tilde{\lambda}}) g_{\tilde{\nu}}{}^{[\mu}(x^{\tilde{\lambda}}, z^{\lambda})\sigma^{\nu]}(x^{\tilde{\lambda}}, z^{\lambda}), \tag{3.21b}$$

where $g_{\mu\tilde{\nu}} \equiv -\sigma_{\mu\tilde{\nu}}$ is the parallel propagator bitensor in flat spacetime.

We modify the Dixon-Harte framework in the following ways.

- We specialize the parameter τ to be the proper time.
- We dispense with the unit vector $n^{\alpha}(\tau)$.
- We use the stress energy tensor $T^{\mu\nu}$ of Eq. (3.15) instead of the matter stress energy tensor $T^{\mu\nu}_{M}$.
- We use null hypersurfaces Σ_τ that are generated by the set of future null geodesics starting at worldline point $z^{\alpha}(\tau)$. This family of null hypersurfaces foliates the convex normal neighborhood of the worldline, which covers the entire manifold for the flat spacetime case we consider in this paper.

Our definitions are then

$$P^{\mu}_{B}(\tau) = \int_{\Sigma_\tau} d\Sigma_{\tilde{\mu}}(x^{\tilde{\lambda}}) T^{\tilde{\mu}\tilde{\nu}}(x^{\tilde{\lambda}}) g_{\tilde{\nu}}{}^{\mu}(x^{\tilde{\lambda}}, z^{\lambda}(\tau)), \tag{3.22a}$$

$$S^{\mu\nu}_{B}(\tau) = 2\int_{\Sigma_\tau} d\Sigma_{\tilde{\mu}}(x^{\tilde{\lambda}}) T^{\tilde{\mu}\tilde{\nu}}(x^{\tilde{\lambda}}) g_{\tilde{\nu}}{}^{[\mu}(x^{\tilde{\lambda}}, z^{\lambda})\sigma^{\nu]}(x^{\tilde{\lambda}}, z^{\lambda}), \tag{3.22b}$$

Here the subscript B denotes "bare"; these definitions will be replaced by renormalized momentum and spin in Sec. 3.4.6 below.

The motivations for our choice of foliation of future null cones are as follows. The integrals (3.17) contain a contribution from the stress energy tensor of the self field from Eq. (3.15). That self field, evaluated at a point x on the hypersurface Σ_τ over which one integrates, in turn depends on the body's charge distribution on the past light cone of x. When one uses a spacelike hypersurface Σ_τ, the dependence on the body's charge distribution extends into the distant past, as one takes x further and further out on the spacelike hypersurface. By contrast, for a future null cone, Σ_τ, the dependence on the body's charge distribution is limited to times within a light-crossing time of τ, as illustrated in figure (3.2). In addition, we show in Appendix 3.A that the integrals (3.22) are well defined and finite when the hypersurfaces Σ_τ are chosen to be future null cones.

There are three choices we have alluded to in the above definition of momentum and spin: the worldline $z(\tau)$ (which is fixed by the spin supplementary condition), the choice (3.15) of body stress-energy tensor, and the choice of the hypersurface of integration. As we have argued, not all choices give rise to physically acceptable definitions. Within those that do there is considerable freedom. This freedom corresponds to different ways of describing a given dynamical system. Different choices will give rise to different forms of the laws of motion, but will not change any physical predictions.

We also define the bare rest mass m_B by

$$m_B^2 \equiv -P_B^\mu P_{B\,\mu}. \tag{3.23}$$

We define the 4-velocity in the usual way as $u^\mu(\tau) = dz^\mu/d\tau$, with $u^\mu u_\mu = -1$, and note that

$$P_B^\mu \neq m_B u^\mu, \tag{3.24}$$

beyond leading order.

The definitions (3.22) are valid for any choice of worldline z_τ. To pick out a unique worldline one must specify a spin supplementary condition [181, 182], which takes the generic form

$$S_B^{\mu\nu}(\tau)\omega_\nu = 0, \tag{3.25}$$

where ω_ν is some vector field defined on the worldline. Such a spin supplementary condition defines a center of mass worldline [183] [184]. Our the spin supplementary condition is defined in terms of

a renormalized spin $S^{\mu\nu}$, which we define in Eq. (3.61) below. Our spin supplementary condition is

$$S^{\mu\nu}u_\nu = 0, \tag{3.26}$$

which reduces at leading order in the size and mass of the body to the condition (3.25) with $\omega_\nu = u_\nu$.

3.2.3 Electromagnetic multipole moments

We now turn to a discussion of electromagnetic multipole moments. We define the total (conserved) bare charge q_B, charge moment $\mathcal{J}^\mu_B$, dipole $Q^{\mu\nu}_B$, and quadrupole $Q^{\mu\nu\rho}_B$ of the body to be

$$q_B(\tau) = q_B = \int_{\Sigma_\tau} d\Sigma_{\tilde\nu} j^{\tilde\nu}, \tag{3.27a}$$

$$\mathcal{J}^\mu_B(\tau) = \int_{\Sigma_\tau} d\Sigma_{\tilde\nu} g^{\tilde\nu}{}_\lambda u^\lambda j^{\tilde\mu} g_{\tilde\mu}{}^\mu, \tag{3.27b}$$

$$Q^{\mu\nu}_B(\tau) = -\int_{\Sigma_\tau} d\Sigma_{\tilde\nu} g^{\tilde\nu}{}_\lambda u^\lambda j^{\tilde\mu} g_{\tilde\mu}{}^\mu \sigma^\nu, \tag{3.27c}$$

$$Q^{\mu\nu\rho}_B(\tau) = \int_{\Sigma_\tau} d\Sigma_{\tilde\nu} g^{\tilde\nu}{}_\lambda u^\lambda j^{\tilde\mu} g_{\tilde\mu}{}^\mu \sigma^\nu \sigma^\rho. \tag{3.27d}$$

In these expressions, the arguments of all the bitensors $g^{\tilde\nu}{}_\lambda$, σ^ν, etc. are $(x, z(\tau))$, while the argument of $j^{\tilde\mu}$ is (x). The definition (3.27c) has a minus sign due to the properties of Synge's worldfunction ($g_\mu{}^{\tilde\nu}\sigma^\mu = -\sigma^{\tilde\nu}$).

For the Dixon moments [182] defined in terms of a spacelike hypersurface generated by geodesics orthogonal to $n^\mu(\tau)$, the bitensor $\sigma^\mu(x, z(\tau))$ is orthogonal to $n^\mu(\tau)$ for all x in Σ_τ, and hence all of the charge moments are orthogonal to n^μ in all indices following the first index:

$$Q^{\mu\nu}_D n_\mu = Q^{\mu\nu\rho}_D n_\mu = Q^{\mu\nu\rho}_D n_\rho = 0. \tag{3.28}$$

Since we integrate over future-directed null cones, there is no such orthogonality condition for our moments (3.27). In addition, our dipole (3.27c) contains both a symmetric and an antisymmetric part, unlike the case for the standard definition which includes an explicit antisymmetrization.

The number of independent components of the electromagnetic dipole (3.27c) and quadrupole (3.27d) are nominally 16 and 40, respectively. When charge conservation is imposed in Sec. 3.6.1, we shall see that these reduce to 10 and 22. However, these are still larger than the number of degrees of freedom for the standard definitions of the electromagnetic dipole and quadrupole, which

are 6 and 14. Our bare electromagnetic moments (3.27) are convenient for our derivation in Sec. 3.6. However, we shall express our final results for the equations of motion in terms of a set of renormalized, projected moments, defined in Sec. 3.4.6, which have the standard number of degrees of freedom.

3.2.4 Scalar multipole moments

For the scalar case, we define an analogous set of bare moments, based on integrals over the scalar source ρ,

$$q_{SB}(\tau) = \int_{\Sigma_\tau} d\Sigma_{\tilde{\nu}} u^{\tilde{\nu}} \rho, \tag{3.29a}$$

$$Q^{\mu}_{SB}(\tau) = -\int_{\Sigma_\tau} d\Sigma_{\tilde{\nu}} g^{\tilde{\nu}}{}_{\lambda} u^{\lambda} \rho \sigma^{\mu}, \tag{3.29b}$$

$$Q^{\mu\nu}_{SB}(\tau) = \int_{\Sigma_\tau} d\Sigma_{\tilde{\nu}} g^{\tilde{\nu}}{}_{\lambda} u^{\lambda} \rho \sigma^{\mu} \sigma^{\nu}. \tag{3.29c}$$

All other details regarding the absence of an orthogonality condition, and the comparison to standard multipoles are similar to those for the electromagnetic multipoles. Here the subscript S denotes "scalar" and B denotes "bare".

The multipole moments (3.27) and (3.29) that we are defining are non-standard. However, they contain the same information as standard multipole moments which are defined in terms of integrals over spacelike hypersurfaces. Some insight into the relation between the two sets of moments can be obtained by considering the leading order expansion for Φ in terms of its source ρ in a Lorentz frame (t, x^i):

$$\Phi(t, r, n^i) = \frac{1}{r} \int d^3y \rho(t - r + n \cdot y, y) + \mathcal{O}\left(\frac{1}{r^2}\right), \tag{3.30}$$

where

$$r = |x| \qquad \text{and} \qquad n^i = \frac{x^i}{r}. \tag{3.31}$$

Taylor expanding the density about the retarded time $t - r$ gives the usual multipole expression

$$\Phi(t, r, n^i) = \frac{1}{r} \sum_{k=0}^{\infty} \left[\frac{1}{k!} n^{i_1} \dots n^{i_k} \int d^3y y^{i_1} \dots y^{i_k} \rho^{(k)}(t - r, y) \right] + \mathcal{O}\left(\frac{1}{r^2}\right), \tag{3.32}$$

where $\rho^{(k)}$ denotes the k^{th} time derivative. Taylor expanding instead about $r - t + y$ yields

$$\Phi(t, r, n^i) = \frac{1}{r} \sum_{k=0}^{\infty} \left[\frac{1}{k!} \int d^3y (n^i y^i - y)^k \rho^{(k)}(t - r + y, y) \right] + \mathcal{O}\left(\frac{1}{r^2}\right), \tag{3.33}$$

which now involves integral over the future null cones. The integrals that appear in (3.33) are precisely time derivatives of our nonstandard multipoles (3.29)

3.3 Non-perturbative equations of motion

This paper focuses primarily on a perturbative expansion of the self force. It is informative, though, to consider the extent to which exact computations can be used to determine radiation-reaction effects. In this section, we derive an exact law of motion for extended bodies, which is used indirectly in our derivation in the remainder of the paper. Our exact law is a modification of an exact law of motion due to Harte [174, 181], which we review. We use Harte's result to perform a consistency check of our results in Sec. 3.6 below.

3.3.1 Equation of motion for bare momentum

First, we define a generalized momentum $P_\tau(\vec{\xi})$ as a linear map on vector fields $\xi^{\tilde{\mu}}$ via

$$P_\tau(\vec{\xi}) = \int_{\Sigma_\tau} T^{\tilde{\mu}\tilde{\nu}} \xi_{\tilde{\mu}} d\Sigma_{\tilde{\nu}}. \tag{3.34}$$

Here, as before, we choose the surface Σ_τ of integration to be future-directed null cones. When we specialize $\vec{\xi}$ to be a Killing vector field $\xi_{\tilde{\mu}} = g_{\tilde{\mu}}{}^{\mu}$ or $\xi_{\tilde{\mu}} = 2g_{\tilde{\mu}}{}^{[\sigma}\sigma^{\nu]}$, the resulting quantities (3.34) yields the definitions (3.22) of linear momentum and spin [174].

To compute the time derivative of this generalized momentum, we use the general identity [174]

$$\frac{d}{d\tau}\int_{\Sigma_\tau} v^{\tilde{\mu}} d\Sigma_{\tilde{\mu}} = \int_{\Sigma_\tau} \nabla_{\tilde{\mu}} v^{\tilde{\mu}} m^{\tilde{\lambda}} d\Sigma_{\tilde{\lambda}} + \int_{\partial\Sigma_\tau} v^{\tilde{\mu}} m^{\tilde{\lambda}} dS_{\tilde{\mu}\tilde{\lambda}}, \tag{3.35}$$

valid for any foliation Σ_τ and any vector field $v^{\tilde{\mu}}$. Here $m^{\tilde{\mu}}$ is any vector field that satisfies $m^{\tilde{\lambda}}(d\tau)_{\tilde{\lambda}} = 1$, $dS_{\tilde{\mu}\tilde{\lambda}} = dS_{[\tilde{\mu}\tilde{\lambda}]}$ is the surface area element, and the second term of the right hand side should be interpreted as a limit of integrals over the boundaries of finite regions of Σ_τ. Applying this identity with $v^{\tilde{\mu}} = T^{\tilde{\mu}\tilde{\nu}}\xi_{\tilde{\nu}}$ gives

$$\frac{d}{d\tau}P_\tau(\vec{\xi}) = \int_{\Sigma_\tau} \nabla_{\tilde{\mu}} T^{\tilde{\mu}\tilde{\nu}} \xi_{\tilde{\nu}} m^{\tilde{\lambda}} d\Sigma_{\tilde{\lambda}} + \frac{1}{2}\int_{\Sigma_\tau} T^{\tilde{\mu}\tilde{\nu}} (\mathcal{L}_\xi g)_{\tilde{\mu}\tilde{\nu}} m^{\tilde{\lambda}} d\Sigma_{\tilde{\lambda}} - \int_{\partial\Sigma_\tau} T^{(\mathrm{self})\tilde{\nu}\tilde{\mu}} \xi_{\tilde{\mu}} m^{\tilde{\lambda}} dS_{\tilde{\nu}\tilde{\lambda}}, \tag{3.36}$$

In the last term, we've removed the matter contribution to the stress energy tensor, since it has compact spatial support and so does not contribute to the boundary integral in the asymptotic limit.

Using Eq. (3.16a) we can rewrite the first term of (3.36) in terms of the external field. Specializing to Killing vector fields, for which the second term vanishes, gives

$$\frac{d}{d\tau}P_\tau(\vec{\xi}) = \int_{\Sigma_\tau}\left(F^{(\text{ext})\tilde{\mu}\tilde{\nu}}\xi_{\tilde{\mu}}j_{\tilde{\nu}}\right)m^{\tilde{\lambda}}d\Sigma_{\tilde{\lambda}} - \int_{\partial\Sigma_\tau}T^{(\text{self})\tilde{\mu}\tilde{\nu}}\xi_{\tilde{\nu}}m^{\tilde{\lambda}}dS_{\tilde{\mu}\tilde{\lambda}}. \tag{3.37}$$

To obtain an explicit equation of motion for the worldline, Eq. (3.37) must be supplemented by the spin supplementary condition (3.26) that determines the relationship between the 4-velocity $u^\mu = dz^\mu/d\tau$ of the worldline and the 4-momentum P_B^μ. To incorporate this condition we proceed as follows. First, we write down the following identities that are valid for any choice of vector field P_B^μ along the worldline

$$m_B a^\kappa = a^\kappa\left(m_B + P_B^\mu u_\mu\right) + \mathcal{P}^\mu{}_\lambda D_\tau P_B^\lambda - \mathcal{P}^\kappa{}_\nu D_\tau\left(\mathcal{P}^\nu{}_\lambda P_B^\lambda\right), \tag{3.38a}$$

$$D_\tau m_B = D_\tau\left(m_B + P_B^\mu u_\mu\right) - u_\mu D_\tau P_B^\mu - a_\mu P_B^\mu. \tag{3.38b}$$

Here $D_\tau \equiv u^\mu\nabla_\mu$ is the covariant derivative along the worldline, $a^\kappa = D_\tau u^\kappa$ is the 4-acceleration, and

$$\mathcal{P}^\mu{}_\lambda = \delta^\mu{}_\lambda + u^\mu u_\lambda \tag{3.39}$$

is the projection tensor onto the space of vectors orthogonal to the 4-velocity. The second term in each of Eqs (3.38a), (3.38b) can be obtained from (3.37) with the choice $\xi_{\tilde{\mu}} = g_{\tilde{\mu}}{}^\mu$ and the replacement $d/d\tau \to D_\tau$. For the first and third terms, we use the general identity (3.35) specialized to

$$v^{\tilde{\mu}} = \sigma^\mu T^{\tilde{\mu}\tilde{\lambda}} n_{\tilde{\lambda}}, \tag{3.40}$$

where $n_{\tilde{\lambda}} = -(d\tau)_{\tilde{\lambda}}$ is the null normal to the future null cone Σ_τ. Using $\nabla_{\tilde{\mu}}\sigma^\mu = -g_{\tilde{\mu}}{}^\mu$, Eq. (3.16a), and the identity for any vector field $v^{\tilde{\mu}}$:

$$\int_{\Sigma_\tau} v^{\tilde{\mu}}d\Sigma_{\tilde{\mu}} = -\int_{\Sigma_\tau} v^{\tilde{\mu}}n_{\tilde{\mu}}m^{\tilde{\lambda}}d\Sigma_{\tilde{\lambda}}, \tag{3.41}$$

we obtain an expression for the bare momentum:

$$\begin{aligned} P_B^\mu(\tau) =& D_\tau\int_{\Sigma_\tau}\sigma^\mu T^{\tilde{\mu}\tilde{\lambda}}n_{\tilde{\lambda}}d\Sigma_{\tilde{\mu}} + \int_{\Sigma_\tau}\sigma^\mu\left[F^{(\text{ext})\tilde{\lambda}\tilde{\rho}}j_{\tilde{\rho}} - T^{\tilde{\mu}\tilde{\rho}}\nabla_{\tilde{\mu}}n_{\tilde{\rho}}m^{\tilde{\lambda}}\right]d\Sigma_{\tilde{\lambda}} \\ &- \int_{\partial\Sigma_\tau}\sigma^\mu m^{[\tilde{\lambda}}T^{\tilde{\mu}]\tilde{\rho}}n_{\tilde{\rho}}dS_{\tilde{\mu}\tilde{\lambda}}. \end{aligned} \tag{3.42}$$

Using the method of Appendix 3.A, one can show that the boundary term in (3.42) vanishes when we choose $\vec{m} = \partial/\partial\tau$ in the coordinates constructed in 3.6.1. The expression (3.42) can now be substituted into the right hand sides of Eqs. (3.38a) and (3.38b) to give explicit evolution equations for the worldline $z^\mu(\tau)$ and bare mass $m_B(\tau)$.

In Sec. 3.5.2 below we will describe a limit in which the charge, mass, and size of the body all go to zero. In this limit, the right hand sides of Eqs. (3.37) and (3.42) can be expanded in terms of electromagnetic multipole moments discussed in Sec. 3.2.3, thereby yielding the explicit form of the equation of motion in this limit. This calculation is carried out in Sec. 3.6. Some of our calculations will proceed directly by taking moments of the field equations (3.7) and (3.16), rather than using Eqs. (3.37) and (3.42).

3.3.2 Equation of motion for Harte's momentum

We now describe an alternative non-perturbative equation of motion for the momentum of extended charged bodies in Minkowski spacetime, due to Harte [174]. It is based on Harte's generalized momentum,

$$P_{H\,\tau}(\vec{\xi}) = \int_{\Sigma_\tau} T_M^{\tilde{\mu}\tilde{\nu}} \xi_{\tilde{\mu}} d\Sigma_{\tilde{\nu}} + E_\tau(\vec{\xi}). \tag{3.43}$$

Here the first term coincides with our bare generalized momentum (3.34), but omits the self-field contribution. The second term $E_\tau(\vec{\xi})$ is a kind of self-field contribution, and is given by Eq.(184) of Ref. [174]. It is a double integral over spacetime that is quadratic in the source $j^{\tilde{\mu}}$, involves a Greens function, and depends on the source only at times τ' that are within a light-crossing time of τ. Its explicit form will not be needed in what follows.

Harte's non-perturbative equation of motion is

$$\frac{d}{d\tau} P_{H\,\tau}(\vec{\xi}) = \int_{\Sigma_\tau} d\Sigma_{\tilde{\nu}} m^{\tilde{\nu}} \left(F^{\tilde{\lambda}\tilde{\rho}} - F_S^{\tilde{\lambda}\tilde{\rho}} \right) \xi_{\tilde{\lambda}} j_{\tilde{\rho}}, \tag{3.44}$$

for Killing vectors $\vec{\xi}$, where $F_S^{\tilde{\lambda}\tilde{\rho}}$ is the average of retarded and advanced self-fields. Harte incorporates the spin-supplementary condition by solving explicitly for the relationship between the 4-velocity and momentum with a choice of parameter τ which differs from proper time. We find it more convenient to proceed instead as described above using the general identity (3.38) and choosing τ to be proper time.

We shall make use of Harte's equation (3.44) as a partial consistency check of our results. By subtracting Eqs.(3.36) and (3.44), we obtain

$$\int_{\Sigma_\tau} d\Sigma_{\tilde{\nu}} m^{\tilde{\nu}} F_R^{\tilde{\lambda}\tilde{\rho}} \xi_{\tilde{\lambda}} j_{\tilde{\rho}} + \int_{\partial\Sigma_\tau} T^{(\text{self})\tilde{\mu}\tilde{\nu}} \xi_{\tilde{\nu}} m^{\tilde{\lambda}} dS_{\tilde{\mu}\tilde{\lambda}} = \begin{pmatrix} \text{Some total} \\ \text{time derivative} \end{pmatrix}, \tag{3.45}$$

where $F_R^{\tilde{\lambda}\tilde{\rho}}$ is the radiative self-field, one half the retarded field minus one half the advanced field. We compute the left hand side of explicitly in terms of our multipole expansion and verify that it is a total time derivative at each order in the expansion; see Secs. 3.6.4.2 and 3.6.5.2 below.

3.4 The point particle limit in the electromagnetic case

3.4.1 One parameter families of solutions: the Gralla-Harte-Wald axioms

We will consider a small charged body interacting with an external electromagnetic field. To describe the limit in which the body becomes very small, we consider a one-parameter family of solutions of the field equations for the body, labeled by a dimensionless parameter λ. Following GHW, we impose the following axioms on the family of solutions. The axioms enforce that the mass and charge of the body go to zero as the size goes to zero.

Axiom 1 *There exists a one-parameter family of fields consisting of the Maxwell tensor $F_{\mu\nu}(\lambda, x^\mu)$, the charge current density $j^\mu(\lambda, x^\mu)$, and the stress-energy tensor $T_M^{\mu\nu}(\lambda, x^\mu)$, which satisfy the Maxwell, charge current conservation and stress-energy conservation equations:*

$$\nabla^\nu F_{\mu\nu}(\lambda, x^\mu) = 4\pi j_\nu(\lambda, x^\mu), \tag{3.46a}$$

$$\nabla_{[\mu} F_{\nu\lambda]} = 0, \tag{3.46b}$$

$$\nabla_\mu j^\mu(\lambda, x^\mu) = 0, \tag{3.46c}$$

$$\nabla_\mu T^{\mu\nu}(\lambda, x^\mu) = 0, \tag{3.46d}$$

where $T^{\mu\nu} \equiv T_M^{\mu\nu} + T_F^{\mu\nu}$, and $T_F^{\mu\nu}$ is given by (3.9). These fields are defined on the open interval $0 < \lambda < \lambda_0$, for some λ_0.

Axiom 2 *We assume there exist functions* $\tilde{j}^\mu(\lambda, t, X^i)$ *and* $\tilde{T}^{\mu\nu}_M(\lambda, t, X^i)$ *such that for some global Lorentz frame coordinates* (t, x^i)*:*

$$j^\mu(\lambda, t, x^i) = \lambda^{-2}\tilde{j}^\mu\left(\lambda, t, \frac{x^i - z^i(\lambda, t)}{\lambda}\right), \tag{3.47a}$$

$$T^{\mu\nu}_M(\lambda, t, x^i) = \lambda^{-2}\tilde{T}^{\mu\nu}_M\left(\lambda, t, \frac{x^i - z^i(\lambda, t)}{\lambda}\right), \tag{3.47b}$$

where $\tilde{j}^\mu$ *and* $\tilde{T}^{\mu\nu}_M$ *are jointly smooth all of in their arguments, including at* $\lambda = 0$*, and* $z^i(\lambda, t)$ *is the center-of mass worldline defined by (3.26).*

Axiom 3 *All of the fields* $F_{\mu\nu}$*,* j^μ*, and* $T^M_{\mu\nu}$ *are jointly smooth in* x^μ *and* λ *away from* $\lambda = 0$*. There exists a worldtube* $\mathcal{W}$ *of compact spatial support such that the supports of* $\tilde{j}^\mu$ *and* $\tilde{T}^{\mu\nu}_M$ *lie inside* $\mathcal{W}$ *for all* λ*.*

Axiom 4 *The external field* $F^{(\mathrm{ext})\mu\nu}$ *defined by (3.12) is jointly smooth in* $x^{\tilde{\mu}}$ *and* λ*, including at* $\lambda = 0$*.*

3.4.2 Discussion of and motivation for the axioms

As in GHW, the axioms 1-4 are intended to describe a family of physically reasonable charge current and stress-energy distributions, such that the limit $\lambda \to 0$ represents a pointlike object. At any finite λ, however, the object is nonsingular with smooth (in particular, non-distributional) sources and a finite self field. Our goal is to derive a set of ordinary differential equations that govern the motion of the object in the limit of small λ.

The axioms enforce a limit where the size $\mathcal{L}$ of the body is much smaller than the scale[2] $\mathcal{L}_{\mathrm{ext}}$ of variation of the external field $F^{(\mathrm{ext})\mu\nu}$. Thus, there is a separation of scales

$$\mathcal{L} << \mathcal{L}_{\mathrm{ext}}. \tag{3.48}$$

One can think of the parameter λ in our one parameter family of solutions as being the ratio $\mathcal{L}/\mathcal{L}_{\mathrm{ext}}$, since the size of the body decreases linearly with λ, from Eqs. (3.47a) and (3.47b). As discussed by GHW, a crucial feature of the assumed one-parameter family is that the mass and charge of the body go to zero as $\lambda \to 0$, at the same rate as the size.

[2]This scale can either be the characteristic length over which $F^{(\mathrm{ext})}$ varies, or the characteristic time.

Our axioms are identical to those of GHW except for the status of the worldline. GHW assume the existence of a λ-independent worldline $z^i(t)$ for which a version of (3.47), with $z^i(\lambda, t)$ replaced by $z^i(t)$, is satisfied. By contrast, we define a one-parameter family of worldlines $z^i(\lambda, t)$ according to the general prescription described in Sec. 3.2.2. The two approaches coincide at leading order, but at subleading order the λ-dependent worldline is more convenient.

Axiom 2 appears to violate Lorentz invariance by the choice of a specific Lorentz frame. However, if this assumption is satisfied in some Lorentz frame, it is satisfied in all Lorentz frames, so it does not violate Lorentz invariance.

3.4.3 Consequence of axioms: the near zone and far zone limits

Following GHW, it is instructive to consider two different limits of $\lambda \to 0$ that give complementary descriptions of the interaction of the body with the external field.

The limit $\lambda \to 0$ at fixed rescaled coordinates

$$(T, X^i) \equiv \left(t, \frac{x^i - z^i(t, \lambda)}{\lambda}\right), \tag{3.49}$$

describes the "near zone" limit. It describes what would be measured by observers at distances from the object of order the object's size $\mathcal{L}$. In this limit, points with fixed global Lorentzian coordinates x^i become more and more distant as $\lambda \to 0$. The lengthscale $\mathcal{L}_{\text{ext}}$ of the external field goes to infinity, while the size $\mathcal{L}$ of the body remains finite.

The limit $\lambda \to 0$ at fixed (t, x^i) describes the "far zone" limit. It describes what would be measured by observers at distances from the object of order $\mathcal{L}_{\text{ext}}$. In this limit, points at fixed rescaled coordinates (T, X^i) approach the worldline $x^i = z^i(0, t)$ as $\lambda \to 0$. In particular, the object's size $\mathcal{L} \to 0$ as $\lambda \to 0$ at fixed (t, x^i).

The GHW axiom approach is closely related to the matched asymptotics method often used in gravitational calculations [160,163,185–187]. The 'near zone' expressions are analogous to an expansion in positive powers of the radial coordinate, valid near the body, and the 'far zone' expressions are analogous to the expansions approximating the body as a pointlike source.

We now discuss the limiting behavior of the self-field as $\lambda \to 0$. The assumptions of subsection 3.2.1 do not demand smoothness of the matter fields j^μ and $T^{\mu\nu}$ in λ at $\lambda = 0$. As shown by GHW,

it follows from axioms 1-4 that the limits $\lambda \to 0$ of the matter fields j^μ and $T^{\mu\nu}$ exist as distributions. This result reflects the desired "point particle" nature of the $\lambda \to 0$ limit of the body. However, axiom 4 demands that in the limit $\lambda \to 0$, the external field remains smooth in the coordinates x^i. This ensures that the external field possesses a well-defined value at the worldline, even in the point particle limit.

The limiting behavior of the self field is derived in the appendix of [159], and can be described as follows. There exists a function $\tilde{F}^{(\mathrm{self})\mu\nu}$, which is jointly smooth in its arguments, including at $\lambda = 0$, such that

$$F^{(\mathrm{self})\mu\nu}(\lambda, t, x^i) = \lambda^{-1}\tilde{F}^{(\mathrm{self})\mu\nu}(\lambda, t, X^i). \tag{3.50}$$

We define a tilded version of the full electromagnetic field $F^{\mu\nu}(\lambda, t, x^i)$, by

$$\tilde{F}^{\mu\nu}(\lambda, t, X^i) = \lambda F^{\mu\nu}\left[\lambda, t, z^i(t, \lambda) + \lambda X^i\right]. \tag{3.51}$$

It follows from (3.50) that this full field can be written as

$$\tilde{F}^{\mu\nu}(\lambda, t, X^i) = \tilde{F}^{(\mathrm{self})\mu\nu}(\lambda, t, X^i) + \lambda F^{(\mathrm{ext})\mu\nu}(\lambda, t, z^i + \lambda X^i), \tag{3.52}$$

so as $\lambda \to 0$ at fixed X^i, $\tilde{F}^{\mu\nu} \to \tilde{F}^{(\mathrm{self})\,\mu\nu}$. It also follows for (3.50) and (3.9) that the stress-energy tensor (3.15) obeys an axiom of the form (3.47b)

$$T^{\mu\nu}(\lambda, t, x^i) = \lambda^{-2}\tilde{T}^{\mu\nu}\left(\lambda, t, \frac{x^i - z^i(\lambda, t)}{\lambda}\right), \tag{3.53}$$

where the right hand side is a smooth function of its arguments.

3.4.4 Limiting behavior of body parameters

We next specialize the general definitions (3.27) of electromagnetic multipole moments to the one-parameter family of charge currents. We find from Eq.(3.47b) that

$$q_B(\lambda) = \lambda\tilde{q}(\lambda), \tag{3.54a}$$

$$\mathcal{J}^\mu_B(\tau, \lambda) = \lambda\tilde{\mathcal{J}}^\mu(\tau, \lambda), \tag{3.54b}$$

$$Q^{\mu\nu}_B(\tau, \lambda) = \lambda^2\tilde{Q}^{\mu\nu}(\tau, \lambda), \tag{3.54c}$$

$$Q^{\mu\nu\lambda}_B(\tau, \lambda) = \lambda^3\tilde{Q}^{\mu\nu}(\tau, \lambda), \tag{3.54d}$$

where the rescaled moments $\tilde{q}, \tilde{\mathcal{J}}^\mu, \tilde{Q}^{\mu\nu}$, and $\tilde{Q}^{\mu\nu\lambda}$ have Taylor expansions about $\lambda = 0$ that start at $\mathcal{O}(\lambda^0)$, for example

$$\tilde{q}(\lambda) = \tilde{q}^{(0)} + \lambda\tilde{q}^{(1)} + \dots . \tag{3.55}$$

The result (3.54) is one of the principal benefits of using the one-parameter family of solutions: in the limit $\lambda \to 0$, successively higher multipoles are suppressed by a higher and higher power of λ. Hence, the limit enforces a multipole expansion.

Similar results apply to the 4-momentum P^μ_B (3.22a) and spin $S^{\mu\nu}_B$ (3.22b), which can be written as

$$P^\mu_B(\tau, \lambda) = \lambda\tilde{P}^\mu(\tau, \lambda), \tag{3.56a}$$

$$S^{\mu\nu}_B(\tau, \lambda) = \lambda^2\tilde{S}^{\mu\nu}(\tau, \lambda), \tag{3.56b}$$

where $\tilde{P}^\mu$ and $\tilde{S}^{\mu\nu}$ have nonzero limits as $\lambda \to 0$. We define a rescaled mass in terms of the rescaled momentum $\tilde{P}^\mu$,

$$\tilde{m}^2 = -\tilde{P}_\mu\tilde{P}^\mu, \tag{3.57}$$

which satisfies $\lambda\tilde{m} = m_B$, and has a finite, non-zero value in the limit $\lambda \to 0$.

3.4.5 Axioms in the scalar case

We use a set of assumptions closely related to axioms 1-4 for the scalar self force derivation. We replace the charge current j^μ with the charge density ρ, the field strength $F_{\mu\nu}$ with the first derivative of the scalar field $\Phi_{;\mu}$, and Maxwell's equations (3.7) with the Klein-Gordon wave equation (3.8).

The scalar charge moments (3.29) can be written as

$$q_{SB}(\lambda) = \lambda\tilde{q}_S(\lambda), \tag{3.58a}$$

$$Q^\mu_{SB}(\tau, \lambda) = \lambda^2\tilde{Q}^\mu_S(\tau, \lambda), \tag{3.58b}$$

$$Q^{\mu\nu}_{SB}(\tau, \lambda) = \lambda^3\tilde{Q}^{\mu\nu}_S(\tau, \lambda), \tag{3.58c}$$

where $\tilde{q}_S$, $\tilde{Q}^\mu_S$, and $\tilde{Q}^{\mu\nu}_S$ have finite, non-zero limits as $\lambda \to 0$, just as for the electromagnetic moments above.

3.4.6 Renormalized projected body parameters

In this section we define a set of renormalized and projected body parameters - momentum, angular momentum and electromagnetic moments - that have a number of desirable properties:

- The final equation of motion is simpler when expressed in terms of these body parameters rather than the original (bare) body parameters.
- The projected parameters have the conventional number of independent degrees of freedom (6 for electromagnetic dipole, 14 for quadrupole), unlike our original definitions (3.27) which had 10 degrees of freedom for the dipole and 22 for the quadrupole.
- The renormalizations are chosen such that the final equations of motion depend only on the renormalized projected parameters.

Our definitions of renormalized projected body parameters are perturbative and are limited to the context of the one-parameter family of solutions. It would be interesting to find more general, non-perturbative definitions that reduce to these definitions in the $\lambda \to 0$ limit. We have been unable to do so. We do note that our renormalized projected parameters are not all obtained at second order by taking the $\lambda \to 0$ limit of Harte's non-perturbative definitions specialized to a spacelike foliation Σ_τ. We therefore expect that such a procedure will not hold in general, and merely define the renormalized, projected moments perturbatively.

The renormalized mass is given by

$$\begin{aligned} m = &- \tilde{P}^\mu u_\mu - \lambda u_\mu F^{(\mathrm{ext})\mu}{}_\nu \tilde{Q}^{\nu\lambda} u_\lambda - \tfrac{2}{3}\lambda^2 \tilde{q} a_\mu D_\tau \left(\mathcal{P}^\mu{}_\nu \tilde{Q}^{\lambda\nu} u_\lambda \right) \\ &- \lambda^2 u_\mu F^{(\mathrm{ext})\mu}{}_{\nu;\lambda} \mathcal{P}^\lambda{}_\eta \tilde{Q}^{\nu\eta\sigma} u_\sigma + \lambda^2 u_\mu F^{(\mathrm{ext})\mu}{}_\nu \tilde{Q}^{\nu\lambda\eta} a_\lambda u_\eta + \mathcal{O}(\lambda^3), \end{aligned} \tag{3.59}$$

where u^μ is the 4-velocity and a^μ the 4-acceleration of the worldline, $\mathcal{P}_\mu{}^\nu = \delta_\mu{}^\nu + u_\mu u^\nu$ is the projection tensor, and $D_\tau = u^\mu \nabla_\mu$. The rescaled electromagnetic dipole $\tilde{Q}^{\mu\lambda}$ and quadrupole $\tilde{Q}^{\mu\nu\lambda}$ which appear here are defined in Eq. (3.54).

Note that $\tilde{P}^\mu u_\mu = -\tilde{m} + \mathcal{O}(\lambda)$, so m and $\tilde{m}$ coincide to leading order. In the limit $\lambda \to 0$ the renormalized mass can be expanded as

$$m(\lambda) = m^{(0)} + \lambda m^{(1)} + \lambda^2 m^{(2)} + \ldots, \tag{3.60}$$

Bare moments	Rescaled bare moments	Renormalized projected moments
P_B^μ (3.22a) : 4	$\tilde{P}$ (3.56a) : 4	not required
m_B (3.23) : 1	$\tilde{m}$ (3.57) : 1	m (3.59) : 1
$S_B^{\mu\nu}$ (3.22b) : 3	$\tilde{S}^{\mu\nu}$ (3.56b) : 3	$S^{\mu\nu}$ (3.61) : 3
q_B (3.27a) : 1	$\tilde{q}$ (3.54a) : 1	q (3.62) : 1
$\mathcal{J}_B^\mu$ (3.27b) : 0	$\tilde{\mathcal{J}}^\mu$ (3.54b) : 0	not required
$Q_B^{\mu\nu}$ (3.27c) : 10	$\tilde{Q}^{\mu\nu}$ (3.54c) : 10	$Q^{\mu\nu}$ (3.63) : 6
$Q_B^{\mu\nu\lambda}$ (3.27d) : 22	$\tilde{Q}^{\mu\nu\lambda}$ (3.54d) : 22	$Q^{\mu\nu\lambda}$ (3.66) : 14

Table 3.1: A summary of the various body parameters we have defined. Each cell lists the symbol for the quantity, the number of the equation in which the quantity is defined, and the number of independent components in the quantity after the charge conservation and the spin supplementary condition have been imposed.

where the coefficients $m^{(0)}$,$m^{(1)}$, etc are independent of λ and $m^{(0)} \neq 0$.

We do not define a renormalized momentum since the momentum is eliminated in the final equation of motion.

The renormalized spin is

$$\begin{aligned} S^{\mu\nu} =& \tilde{S}^{\mu\nu} + 2\lambda F^{(\text{ext})[\mu|}{}_\lambda \tilde{Q}^{\lambda|\nu]\rho} u_\rho + \tfrac{2}{3}\lambda\tilde{q}\mathcal{P}^{[\mu}{}_\lambda u^{\nu]}\tilde{Q}^{\lambda\rho}a_\rho \\ &+ \lambda\mathcal{P}^{[\mu}{}_\lambda\mathcal{P}^{\nu]}{}_\rho\left(\tfrac{2}{3}\tilde{q}D_\tau\tilde{Q}^{\lambda\rho} + \tfrac{4}{3}\tilde{q}u_\eta\tilde{Q}^{\eta\lambda}a^\rho + \tfrac{2}{3}\tilde{q}\tilde{Q}^{\lambda\eta}u_\eta a^\rho\right) + \mathcal{O}(\lambda^2). \end{aligned} \tag{3.61}$$

This also can be expanded in powers of λ with a leading term which is non-zero.

The charge is conserved so requires no renormalization,

$$q = \tilde{q}. \tag{3.62}$$

The renormalized, projected electromagnetic dipole is

$$Q^{\mu\nu} = \left(\tilde{Q}^{\mu\nu} + \lambda u_\sigma D_\tau\left(\tilde{Q}^{\mu\nu\sigma}\right)\right)\mathcal{P}_\nu{}^\kappa + \mathcal{O}(\lambda^2), \tag{3.63}$$

Note that this dipole is orthogonal to the 4-velocity on its second index, unlike the bare dipole. We can expand $Q^{\mu\nu}$ as

$$Q^{\mu\nu} = Q^{(0)\mu\nu} + \lambda Q^{(1)\mu\nu} + \mathcal{O}(\lambda^2). \tag{3.64}$$

Charge conservation [Eq. (3.102c) below with $m = 2$ and $N = 2$] enforces that the spatial components of the leading order term are antisymmetric,

$$Q^{(0)\mu\nu}\mathcal{P}_\mu{}^{(\lambda}\mathcal{P}_\nu{}^{\eta)} = 0. \tag{3.65}$$

At higher order, the quantity $Q^{(1)\mu\nu}\mathcal{P}_\mu{}^{(\lambda}\mathcal{P}_\nu{}^{\lambda)}$ can be computed from the time derivative of the electric quadrupole and the corresponding subleading charge conservation [Eq. (3.102c), order $\mathcal{O}(\lambda)$, with $m = 2$ and $N = 2$]. Hence, the dipole (3.63) has 6 independent components.

We note that if we replace the future null cone Σ_τ in the definitions (3.27) of electromagnetic moments with a spacelike hypersurface orthogonal to the 4-velocity, then the same final result would be obtained by taking the expression (3.63) but omitting the correction term.

The renormalized, projected quadrupole is

$$Q^{\mu\lambda\eta} = \mathcal{P}^\lambda{}_\nu\mathcal{P}^\eta{}_\sigma\tilde{Q}^{\mu\nu\sigma} + \mathcal{O}(\lambda). \tag{3.66}$$

This tensor is orthogonal to the 4-velocity in its second two indices. The completely symmetric part of the spatial projection of this quadrupole vanishes to leading order

$$Q^{\mu\nu\sigma}\mathcal{P}_\mu{}^{(\lambda}\mathcal{P}_\nu{}^{\eta}\mathcal{P}_\sigma{}^{\rho)} = \mathcal{O}(\lambda), \tag{3.67}$$

from Eq.(3.102c) below with $m = 3$, $N = 3$. It follows that the leading order renormalized quadrupole has the standard number of independent components (6 electric and 8 magnetic).

The notations for and properties of the various body parameters we have defined are summarized in Table 3.1.

3.5 Summary of results: electromagnetic laws of motion

3.5.1 Preamble: domain of validity of self force equations

The classic Abraham-Lorentz-Dirac radiation-reaction equation,

$$a^\nu = \frac{q}{m}F^{\nu\mu}u_\mu + \frac{2}{3}\frac{q^2}{m}\mathcal{P}^\nu{}_\mu\dot{a}^\mu, \tag{3.68}$$

is a third-order differential equation which possesses transparently nonphysical runaway solutions.

As pointed out by GHW, (3.68) is valid only in the regime $q^2\dot{a}/ma \equiv \epsilon \ll 1$, and the equation has errors of order $\epsilon^2 a$. The runaway solutions possess a rapidly growing acceleration, and violate the assumption $\epsilon \ll 1$. When $\epsilon \gtrsim 1$, the perturbative differential equation (3.68) is no longer a good approximation.

The reduction of order procedure provides a method of deriving from Eq. (3.68) an equation which is equally accurate but which is second order in time and which does not have runaway solutions [188–192]. Substituting the expression for the acceleration given by the first term in (3.68) into the second term modifies the equation by a term which is no larger than the pre-existing error terms. The resulting reduced-order equation is

$$a^\sigma = \frac{q}{m} F^{\sigma\mu} u_\mu + \frac{2}{3}\frac{q^3}{m^2} \mathcal{P}^\sigma{}_\rho \left(F^{\rho\mu}{}_{;\nu} u_\mu u^\nu + F^{\rho\mu} F_{\mu\nu} u^\nu \right) + \mathcal{O}(q^5). \tag{3.69}$$

Our final results (3.70) are expressed as an expansion in powers of λ, a parameter which is proportional to the charge q, also the mass m, and here also to q^2/m. We do not perform a reduction of order in our results for brevity. (except the point particle case discussed in Sec. 3.5.4 below). However, we emphasize that our results should be interpreted in terms of their reduced-order counterparts.

3.5.2 | Laws of motion - general self force and center of mass evolution

We present in this section the results for the electromagnetic case. The scalar results are derived in much the same way, and can be found in the appendix 3.B.

The evolution of the body's worldline $z^\mu(\tau)$ and rest mass to second order in λ are given by

$$ma^\mu = f^{(0)\mu} + \lambda f^{(1)\mu} + \lambda^2 f^{(2)\mu} + \mathcal{O}(\lambda^3), \tag{3.70a}$$

$$D_\tau m = \lambda \mathcal{F}^{(1)} + \lambda^2 \mathcal{F}^{(2)} + \mathcal{O}(\lambda^3), \tag{3.70b}$$

where a^μ is the acceleration of the worldline and m is the renormalized mass (3.59). Here $f^{(0)\mu}$ is the Lorentz force, $f^{(1)\mu}$ and $\mathcal{F}^{(1)}$ are the first order GHW results, and $f^{(2)}$ and $\mathcal{F}^{(2)}$ are the new second-order results presented here. Explicit expressions for all these quantities are given in this section and the derivations are given in Sec.3.6 below

We refer to Eqs. (3.70) as 'laws' of motion, instead of equations of motion, as they require additional information about the body's electromagnetic multipoles their time dependence to fully determine the motion. The requisite additional equations parameterize the evolution of the internal degrees of freedom of the body.

At leading order we have the Lorentz force and mass conservation

$$f^{(0)\mu} = qF^{(\text{ext})\mu\lambda}u_\lambda, \tag{3.71a}$$

$$D_\tau m = \mathcal{O}(\lambda). \tag{3.71b}$$

At subleading order we have,

$$\begin{aligned} f^{(1)\kappa} =& \mathcal{P}^\kappa{}_\nu \Big[F^{(\text{ext})\nu}{}_{\mu;\lambda} Q^{\mu\lambda} + \tfrac{2}{3} q D_\tau a^\nu + D_\tau \left(a_\mu S^{\nu\mu}\right) \\ &+ F^{(\text{ext})\nu}{}_\mu D_\tau Q^{[\mu\lambda]} u_\lambda - D_\tau \left(u_\mu F^{(\text{ext})\mu}{}_\lambda Q^{\lambda\nu} \right) \Big], \end{aligned} \tag{3.72a}$$

$$\mathcal{F}^{(1)} = - u_\mu F^{(\text{ext})\mu}{}_{\nu;\lambda} Q^{\nu\lambda} - u_\nu F^{(\text{ext})\nu}{}_\mu D_\tau \left(Q^{\mu\lambda} \right) u_\lambda - 2 u_\mu F^{(\text{ext})\mu}{}_\nu Q^{\nu\lambda} a_\lambda. \tag{3.72b}$$

Here the body's charge q, electromagnetic dipole $Q^{\mu\nu}$, and spin $S^{\mu\nu}$ are the renormalized versions (3.62), (3.63), and (3.61).

To facilitate comparison of the results with those of GHW, we define an antisymmetric dipole $Q_A^{\mu\nu}$ by

$$Q_A{}^{\mu\nu} \mathcal{P}_\nu{}^\lambda = Q^{\mu\lambda}, \tag{3.73a}$$

$$Q_A{}^{\mu\nu} u_\nu = - u_\nu Q^{\nu\mu}, \tag{3.73b}$$

for which $Q_A{}^{(\mu\nu)} = 0$. Eliminating $Q^{\mu\nu}$ in terms of $Q_A^{\mu\nu}$, and we find

$$f^{(1)\kappa} = \mathcal{P}^\kappa{}_\nu \left[F^{(\text{ext})\nu}{}_{\mu;\lambda} Q_A{}^{\mu\lambda} + \tfrac{2}{3} q D_\tau a^\nu + D_\tau \left(a_\mu S^{\nu\mu}\right) + 2 D_\tau \left(u_\mu F^{(\text{ext})\lambda[\nu} Q_A{}^{\mu]}{}_\lambda \right) \right], \tag{3.74a}$$

$$\mathcal{F}^{(1)} = - u_\mu F^{(\text{ext})\mu}{}_{\nu;\lambda} Q_A{}^{\nu\lambda} - D_\tau \left(F^{(\text{ext})\nu}{}_\mu Q_A{}^{\mu\lambda} \right) u_\nu u_\lambda - 2 u_\mu F^{(\text{ext})\mu}{}_\nu Q_A{}^{\nu\lambda} a_\lambda, \tag{3.74b}$$

which agrees with the results of GHW. The third term in the mass evolution (3.74b) does not appear in GHW, however it gives only a $\mathcal{O}(\lambda^2)$ contribution when reduction of order is applied. We retain this term since we will be working to $\mathcal{O}(\lambda^2)$.

As noted in GHW, the first and second terms in the acceleration equation (3.74a) are the monopole self force usually derived from the radiative self field, and the direct interactions with the

external field. The final two terms in (3.74a) are terms that are not usually derived in elementary treatments of electrodynamics.

The second order results can be decomposed into monopole, dipole, and quadrupole contributions:

$$f^{(2)\mu} = f^{(2)\mu}_{\text{point}} + f^{(2)\mu}_{\text{dipole}} + f^{(2)\mu}_{\text{quadrupole}}, \tag{3.75a}$$

$$\mathcal{F}^{(2)} = \mathcal{F}^{(2)}_{\text{point}} + \mathcal{F}^{(2)}_{\text{dipole}} + \mathcal{F}^{(2)}_{\text{quadrupole}}. \tag{3.75b}$$

We have

$$f^{(2)\mu}_{\text{point}} = 0, \tag{3.76a}$$

$$\mathcal{F}^{(2)}_{\text{point}} = 0, \tag{3.76b}$$

so there are no new point particle terms at second order. We note, however, that monopole terms at $\mathcal{O}(\lambda^2)$ would be generated if one expands out the body parameters in a power series in λ, as in Eq. (3.64) above, and also would be generated by the reduction of order procedure, c.f. Sec.3.5.4 below. The explicit, new, dipole and quadrupole contribution to the self force are

$$\begin{aligned} f^{(2)\mu}_{\text{dipole}} = \mathcal{P}^{\sigma}{}_{\kappa}\Big[& -\tfrac{1}{3}qa_\mu a^\mu a_\nu Q^{\nu\kappa} + qa^\kappa D_\tau a^\mu \mathcal{P}_{\mu\nu} Q^{\lambda\nu} u_\lambda + \tfrac{7}{6} qD_\tau a^\kappa a_\mu Q^{\lambda\mu} u_\lambda \\ & - \tfrac{11}{6} qa_\mu D_\tau a^\mu Q^{\nu\kappa} u_\nu + \tfrac{1}{3} qa^\kappa a_\mu D_\tau Q^{\nu\mu} u_\nu - qa_\mu a^\mu D_\tau Q^{\nu\kappa} u_\nu \\ & - \tfrac{2}{3} qD_\tau a_\mu D_\tau Q^{\mu\kappa} - 2qa_\mu D_\tau{}^2 Q^{\mu\kappa} - \tfrac{2}{3} qD_\tau{}^3 Q^{\mu\kappa} u_\mu \Big], \end{aligned} \tag{3.77a}$$

$$\begin{aligned} f^{(2)\mu}_{\text{quadrupole}} = \mathcal{P}^{\sigma}{}_{\kappa}\Big[& \tfrac{1}{2} F^{(\text{ext})\kappa}{}_{\mu;\nu\lambda} Q^{\mu\nu\lambda} - u_\mu D_\tau \left(F^{(\text{ext})\mu}{}_{\nu;\rho} \mathcal{P}^{\nu}{}_{\lambda} Q^{\lambda\kappa\rho} \right) + \tfrac{1}{2} D_\tau{}^2 \left(F^{(\text{ext})\kappa}{}_{\mu} Q^{\mu\rho}{}_{\rho} \right) \\ & - 2u_\mu F^{(\text{ext})\mu}{}_{\lambda;\nu} u^\nu Q^{\lambda\kappa\rho} a_\rho + 2F^{(\text{ext})[\kappa}{}_{\mu;\lambda} Q^{\mu|\nu]\lambda} a_\nu + \tfrac{1}{2} F^{(\text{ext})\kappa}{}_{\mu;\nu} a^\nu Q^{\mu\rho}{}_{\rho} \\ & - \tfrac{1}{2} F^{(\text{ext})\kappa\nu} D_\tau \left(a_\nu u_\mu Q^{\mu\rho}{}_{\rho} \right) - u_\mu F^{(\text{ext})\mu}{}_{\nu} D_\tau \left(Q^{\nu\kappa\lambda} a_\lambda \right) \\ & + a^\kappa u_\mu F^{(\text{ext})\mu}{}_{\nu} D_\tau Q^{\nu\rho}{}_{\rho} - 2a_\nu F^{(\text{ext})(\nu}{}_{\mu} Q^{\mu|\kappa)\lambda} a_\lambda \Big], \end{aligned} \tag{3.77b}$$

and the explicit, new, dipole and quadrupole contributions to the mass evolution are

$$\mathcal{F}^{(2)}_{\text{dipole}} = -\tfrac{1}{3}\tilde{q}a_\mu a^\mu u_\lambda Q_R{}^{\lambda\nu}a_\nu - \tfrac{2}{3}\tilde{q}D_\tau a_\nu \mathcal{P}^\nu{}_\lambda D_\tau\left(u_\mu Q^{\mu\lambda}\right), \tag{3.78a}$$

$$\begin{aligned}\mathcal{F}^{(2)}_{\text{quadrupole}} =& -\tfrac{1}{2}u_\mu F^{(\text{ext})\mu}{}_{\lambda;\nu\rho}Q^{\lambda\nu\rho} - \tfrac{1}{2}u_\mu F^{(\text{ext})\mu}{}_{\lambda;\nu\sigma}u^\nu u^\sigma Q^{\lambda\rho}{}_\rho - 2u_\mu F^{(\text{ext})\mu}{}_{\lambda;\sigma}Q^{\lambda\sigma\nu}a_\nu \\ & - u_\mu F^{(\text{ext})\mu}{}_{\nu;\lambda}a^\lambda Q^{\nu\rho}{}_\rho - \tfrac{1}{2}D_\tau a_\mu F^{(\text{ext})\mu\nu}u_\nu u_\lambda Q^{\lambda\rho}{}_\rho - \tfrac{1}{2}a_\nu F^{(\text{ext})\nu\mu}u_\mu a_\lambda Q^{\lambda\rho}{}_\rho \\ & - \tfrac{1}{2}a_\lambda F^{(\text{ext})\lambda\mu}u_\mu u_\nu D_\tau Q^{\nu\rho}{}_\rho + a_\nu F^{(\text{ext})\nu}{}_\mu D_\tau Q^{\mu\rho}{}_\rho + \tfrac{1}{2}u_\mu F^{(\text{ext})\mu}{}_\lambda D_\tau{}^2 Q^{\lambda\rho}{}_\rho.\end{aligned} \tag{3.78b}$$

3.5.3 | Laws of motion - evolution of spin

Like the self force, the torque may also be written in terms of the renormalized dipole, quadrupole and spin introduced in Sec.3.4.6. The result is

$$\begin{aligned}D_\tau S^{\lambda\rho}\mathcal{P}_\lambda{}^\kappa\mathcal{P}_\rho{}^\sigma =& \mathcal{P}^\kappa{}_\lambda\mathcal{P}^\sigma{}_\rho\Big(2F^{(\text{ext})[\lambda}{}_\mu Q^{\mu|\rho]} + 2\lambda F^{(\text{ext})[\lambda}{}_{\nu;\mu}Q^{\nu\mu|\rho]} \\ & - \tfrac{4}{3}\lambda q D_\tau a^{[\lambda}Q^{\mu|\rho]}u_\mu + 2\lambda F^{(\text{ext})[\lambda}{}_\mu Q^{\mu\nu|\rho]}a_\nu\Big) + \mathcal{O}(\lambda^2).\end{aligned} \tag{3.79}$$

Because of the spin supplementary condition (3.26), this projected version of $D_\tau S^{\lambda\rho}$ is sufficient to determine the entire time derivative. The first term in this torque expression reproduces the GHW result.

3.5.4 | Laws of motion - reduced order point particle limit

In this section, we specialize to monopole bodies, i.e. those with vanishing spin $S^{\mu\nu}$, electromagnetic dipole $Q^{\mu\nu}$, and electromagnetic quadrupole $Q^{\mu\nu\lambda}$. The equations of motion (3.70) then reduce to

$$ma^\mu = \lambda q F^{(\text{ext})\mu\lambda}u_\lambda + \tfrac{2}{3}\lambda^2 q^2 \mathcal{P}^\mu{}_\nu D_\tau a^\nu + \mathcal{O}(\lambda^4), \tag{3.80a}$$

$$D_\tau m = \mathcal{O}(\lambda^3). \tag{3.80b}$$

We now apply a reduction of order to determine the acceleration through $\mathcal{O}(\lambda^2)$ in terms of the external field. The resulting acceleration, given explicitly for the first time, is

$$\begin{aligned}a^\mu =& \frac{q}{m}F^{(\text{ext})\mu\nu}u_\nu + \frac{2q^3}{3m^2}D_\tau F^{(\text{ext})\mu\nu}u_\nu + \frac{2q^4\lambda}{3m^3}\mathcal{P}^\mu{}_\eta F^{(\text{ext})\eta\nu}F^{(\text{ext})}{}_{\nu\sigma}u^\sigma \\ & + \frac{4q^5}{9m^3}\lambda^2 D_\tau{}^2 F^{(\text{ext})\mu\nu}u_\nu + \frac{4q^6}{9m^4}\lambda^2\mathcal{P}^\mu{}_\rho\Big(2D_\tau F^{(\text{ext})\rho\nu}F^{(\text{ext})}{}_{\nu\lambda}u^\lambda + F^{(\text{ext})\rho\nu}D_\tau F^{(\text{ext})}{}_{\nu\lambda}u^\lambda\Big) \\ & + \frac{4q^7}{9m^5}\lambda^2\mathcal{P}^\mu{}_\rho F^{(\text{ext})\rho\nu}\mathcal{P}_{\nu\eta}F^{(\text{ext})\eta\lambda}F^{(\text{ext})}{}_{\lambda\sigma}u^\sigma + \mathcal{O}\left(\lambda^3\right).\end{aligned} \tag{3.81}$$

3.6 Details of derivation

3.6.1 Preliminary definitions and constructions

The derivation is based on the axioms described in sec 3.4.1, which are expressed in some global Lorentz frame coordinates (t, x^i). For the purposes of our derivation, we adopt a retarded body-following coordinate system, motivated by the scaled coordinates (T, X^i) considered in Sec. 3.4.1.

We choose a tetrad at a point on the worldline, $z^\mu(\tau, \lambda)$[3],

$$\{e_{\hat{0}}{}^\mu, e_{\hat{i}}{}^\mu\} \equiv \{u^\mu, e^\mu{}_{\hat{i}}\}, \tag{3.82}$$

which we constrain to be orthonormal:

$$\vec{e}_{\hat{a}} \cdot \vec{e}_{\hat{b}} = \eta_{\hat{a}\hat{b}}. \tag{3.83}$$

We extend this tetrad along the worldline using Fermi-Walker transport

$$\frac{De^\mu{}_{\hat{a}}}{d\tau} = e^\nu_{\hat{a}} \left(u^\mu a_\nu - a^\mu u_\nu\right), \tag{3.84}$$

and extend it off the worldline by parallel transport along generators of future null cones that originate on the worldline.

Tetrad indices are raised and lowered using $\eta_{\hat{a}\hat{b}}$:

$$u^\mu = e^\mu{}_{\hat{0}} = -e^{\mu\hat{0}} \qquad\qquad e^\mu{}_{\hat{j}} = e^{\mu\hat{i}}\delta_{\hat{i}\hat{j}}. \tag{3.85}$$

We next define the retarded Fermi coordinate system $(T, y^{\hat{i}})$ following Poisson [160]. For a given spacelike point $x^{\tilde{\mu}}$, we define $\tau(x^{\tilde{\mu}})$ such that $z^\mu(\tau)$ is the intersection of the past lightcone of $x^{\tilde{\mu}}$ with the worldline, so that

$$\sigma(z^\mu(\tau(x)), x^{\tilde{\mu}}) = 0. \tag{3.86}$$

Surfaces of constant τ are future light cones of points on the worldline. We define the spatial coordinates y^i by

$$y^{\hat{i}} = -\delta^{\hat{i}\hat{j}} e_{\hat{j}}{}^\mu(\tau)\sigma_\mu(z_\tau, x), \tag{3.87}$$

[3]Note that our construction is based on the λ-dependent worldline $z^\mu(\tau, \lambda)$, and not on the fixed, λ-independent worldline $z^\mu(\tau, 0)$.

evaluated at $\tau = \tau(x)$. In these coordinates the metric takes the form [160]

$$ds^2 = -(\varphi^2 - r^2 a^2)d\tau^2 + (\delta_{\hat{i}\hat{j}} - n_{\hat{i}} n_{\hat{j}})dy^{\hat{i}} dy^{\hat{j}} + 2(r a_{\hat{i}} - \varphi n_{\hat{i}})dx^{\hat{i}} d\tau, \tag{3.88}$$

where $r^2 = \delta_{\hat{i}\hat{j}} y^{\hat{i}} y^{\hat{j}}$, $\varphi = 1 + y^{\hat{i}} a_{\hat{i}}$, $n^{\hat{i}} = y^{\hat{i}}/r$. The orthonormal basis in these coordinates is given by

$$\vec{e}_{\hat{0}} = \partial_\tau - r a^{\hat{i}} \partial_{\hat{i}}, \tag{3.89a}$$

$$\vec{e}_{\hat{i}} = \left(\delta_{\hat{i}}^{\hat{j}} + r n_{\hat{i}} a^{\hat{j}}\right) \partial_{\hat{j}} - n_{\hat{i}} \partial_\tau. \tag{3.89b}$$

Next we re-express axiom 2 of Sec. 3.4.1 in terms of these coordinates and the orthonormal basis components of the tensors. From Eq. (3.53), it takes the form

$$T^{\hat{a}\hat{b}}(\lambda, \tau, y^{\hat{i}}) = \lambda^{-2} \tilde{T}^{\hat{a}\hat{b}} \left(\lambda, \tau, y^{\hat{i}}/\lambda\right), \tag{3.90a}$$

$$j^{\hat{a}}(\lambda, \tau, y^{\hat{i}}) = \lambda^{-2} \tilde{j}^{\hat{a}} \left(\lambda, \tau, y^{\hat{i}}/\lambda\right), \tag{3.90b}$$

where the right hand sides are smooth functions of their arguments [distinct from the functions in (3.47a) and (3.53)].

Finally, we can write the rescaled body parameters of Sec. 3.4.4 in terms of the functions $\tilde{T}^{\hat{a}\hat{b}}$ and $j^{\hat{a}}$:

$$\tilde{P}^{\hat{a}} = \int d^3Y \left(\tilde{T}^{\hat{a}\hat{0}} - \tilde{T}^{\hat{a}\hat{i}} n_{\hat{i}}\right), \tag{3.91a}$$

$$\tilde{S}^{\hat{a}\hat{b}} = 2 \int d^3Y R \left(n^{[\hat{a}} \tilde{T}^{\hat{b}]\hat{0}} - n^{[\hat{a}} \tilde{T}^{\hat{b}]\hat{i}} n_{\hat{i}}\right), \tag{3.91b}$$

and

$$\tilde{q} = \int d^3Y (\tilde{j}^{\hat{0}} - \tilde{j}^{\hat{i}} n_{\hat{i}}), \tag{3.92a}$$

$$\tilde{\mathcal{J}}^{\hat{a}} = \int d^3Y \tilde{j}^{\hat{a}}, \tag{3.92b}$$

$$\tilde{Q}^{\hat{a}\hat{b}} = \int d^3Y R \tilde{j}^{\hat{a}} n^{\hat{b}}, \tag{3.92c}$$

$$\tilde{Q}^{\hat{a}\hat{b}\hat{c}} = \int d^3Y R^2 \tilde{j}^{\hat{a}} n^{\hat{a}} n^{\hat{b}}, \tag{3.92d}$$

where $Y^{\hat{i}} = y^{\hat{i}}/\lambda$, $R^2 = \delta_{\hat{i}\hat{j}} Y^{\hat{i}} Y^{\hat{j}}$, and $\vec{n} = \vec{u} + n^{\hat{i}} \vec{e}_{\hat{i}}$. Here the integrals are over surfaces of constant τ, i.e. the future light cones.

3.6.2 Retarded and advanced self-field

In this subsection, we compute the near-zone expansion of the retarded field in terms of the scaled multipoles (3.54) and the retarded coordinates from Sec. 3.6.1. The computation is used in sections 3.6.4-3.6.5.

Consider a field point $x^{\tilde{\mu}}$. Recall that $\tau(x^{\tilde{\mu}})$ denotes the proper time at which the past lightcone of $x^{\tilde{\mu}}$ intersects the wordline $z^{\mu}(\tau)$. We denote by $\mathcal{W}_{-}(x^{\tilde{\mu}})$ the intersection of the interior of the past lightcone of $x^{\tilde{\mu}}$ and the worldtube $\mathcal{W}$ of the body. The retarded, Lorenz-gauge self-field of the body can be written as

$$\begin{aligned} A^{\tilde{\mu}}_{-}(x) &= \int d^4x' \sqrt{-g(x')} G_{-}{}^{\tilde{\mu}}{}_{\nu'}(x,x') j^{\nu'}(x') \\ &= \int_{\mathcal{W}_{-}} d^4x' g^{\tilde{\mu}}{}_{\nu'}(x,x') \delta(\sigma(x,x')) j^{\nu'}(x'), \end{aligned} \tag{3.93}$$

where $G^{\mu}_{-}{}_{\nu}(x,x')$ is the retarded propagator in Lorenz gauge. Here, $g^{\mu}{}_{\nu'}$ is the parallel propagator, and the 1-dimensional delta function $\delta(\sigma(x,x'))$ constrains the integral to the three-surface formed by the past null cone of the field point x.

To relate the right hand side of (3.93) to the bare multipoles (3.27), we wish to write the integral (3.93) as a series of integrals over the future null cone of the intersection point of the center-of-mass worldline (3.26) and the past null cone of $x^{\tilde{\mu}}$, which we will write as $z(\tau)$.

To this end, we write $x^{\tilde{\mu}} = (\tau, y^{\hat{i}})$ and $x'^{\tilde{\mu}'} = (\tau', y^{\hat{i}})$ in the retarded coordinates of Sec. 3.6.1 above. We denote the value of τ' at which σ vanishes as

$$\tau' = \tau + \Delta\tau(\tau, y^{\hat{i}}, y'^{\hat{i}}). \tag{3.94}$$

The δ-function $\delta(\sigma)$ can now be written as

$$\delta(\sigma(x^{\tilde{\mu}}, x'^{\tilde{\nu}'})) = \frac{\delta(\tau' - \tau - \Delta\tau)}{|\sigma_{,\tau}(\tau, y^{\hat{i}}; \tau + \Delta\tau, y'^{\hat{i}})|}. \tag{3.95}$$

Inserting this into Eq. (3.93), using the fact that $|\det(g_{\alpha\beta})| = 1$ in the retarded coordinates, and multiplying by a parallel propagator factor gives

$$A^{\tilde{\mu}}_{-}(\tau, y^{\hat{i}}) g_{\tilde{\mu}}{}^{\mu}(\tau, 0; \tau, y^{\hat{i}}) = \int d^3y' \frac{g^{\mu}{}_{\tilde{\nu}'}(\tau, 0; \tau + \Delta\tau, y'^{\hat{i}}) j^{\tilde{\nu}'}(\tau + \Delta\tau, y'^{\hat{i}})}{|\sigma_{,\tau'}(\tau, y^{\hat{i}}; \tau + \Delta\tau, y'^{\hat{i}})|}. \tag{3.96}$$

We now rewrite this expression in terms of the rescaled spatial coordinates $Y^{\hat{i}} = y^{\hat{i}}/\lambda$, $Y'^{\hat{i}} = y'^{\hat{i}}/\lambda$ and in terms of the tilded version of the charge current from Eq. (3.90). Noting that $\Delta\tau(\tau, \lambda Y^{\hat{i}}, \lambda Y'^{\hat{i}})$ vanishes as $\lambda \to 0$ at fixed $Y^{\hat{i}}$,$Y'^{\hat{i}}$, we write this quantity as

$$\Delta(\tau, \lambda Y^{\hat{i}}, \lambda Y'^{\hat{i}}) = \lambda \widetilde{\Delta\tau}(\tau, Y^{\hat{i}}, Y'^{\hat{i}}, \lambda), \tag{3.97}$$

where $\widetilde{\Delta\tau}$ is finite as $\lambda \to 0$. The result is

$$A^{\tilde{\mu}}_{-}(\tau, Y^{\hat{i}}) = \lambda \int d^3Y' \left[g^{\tilde{\mu}}{}_{\tilde{\nu}'}(\tau, \lambda Y^{\hat{i}}; \tau + \lambda\widetilde{\Delta\tau}, \lambda Y'^{\hat{i}}) \frac{\tilde{j}^{\tilde{\nu}'}(\tau + \lambda\widetilde{\Delta\tau}, Y'^{\hat{i}})}{|\sigma_{,\tau'}(\tau, \lambda Y^{\hat{i}}; \tau + \lambda\widetilde{\Delta\tau}, \lambda Y'^{\hat{i}})|} \right]. \tag{3.98}$$

Finally, we expand the right hand side in powers of λ, and we also take the large $R = |Y|$ limit. Expressing the result in terms of components on the orthonormal tetrad, the retarded field can naturally be expressed in terms of the rescaled electromagnetic moments (3.92)

$$\begin{aligned} A^{\hat{a}}_{-} =& \frac{\tilde{\mathcal{J}}^{\hat{a}}}{R} + \frac{\tilde{Q}^{\hat{a}\hat{j}} n_{\hat{j}}}{R^2} + \lambda a_{\hat{i}} n^{\hat{i}} \frac{\tilde{Q}^{\hat{a}\hat{j}} n_{\hat{j}}}{R} + \lambda(a^{\hat{a}} u_{\hat{b}} - u^{\hat{a}} a_{\hat{b}}) \frac{\tilde{Q}^{\hat{b}\hat{0}} - \tilde{Q}^{\hat{b}\hat{j}} n_{\hat{j}}}{R} - \lambda \frac{\tilde{Q}^{\hat{a}\hat{j}} a_{\hat{j}}}{R} \\ &+ \lambda \frac{\partial_\tau \tilde{Q}^{\hat{a}\hat{j}} n_{\hat{j}}}{R} - \lambda \frac{\partial_\tau \tilde{Q}^{\hat{a}\hat{0}}}{R} + \mathcal{O}\left(\frac{\lambda^n}{R^m}\right), \end{aligned} \tag{3.99}$$

where the omitted terms satisfy $n + m \geq 3$.

We use the result (3.99) to evaluate certain boundary terms at infinity that arise in Sec. 3.6.3 below.

3.6.3 | Moments of the field equations

We next express the fundamental equation (3.16a) and charge current conservation $\nabla_\mu j^\mu = 0$ in terms of the coordinates $(\tau, Y^{\hat{i}})$, using the tilded functions on the right hand sides of (3.90). We use tetrad component of the tensors but write the derivatives in terms of the partial derivatives with respect to the coordinates; this unusual combination is the most convenient for our derivation. The result is:

$$\begin{aligned} \lambda F^{(\mathrm{ext})\hat{k}\hat{i}} \tilde{j}_{\hat{i}} + \lambda F^{(\mathrm{ext})\hat{k}\hat{0}} \tilde{j}_{\hat{0}} =& T^{\hat{k}\hat{j}}{}_{,\hat{j}} + \lambda T^{\hat{k}\hat{0}}{}_{,0} - \lambda n_{\hat{i}} T^{\hat{k}\hat{i}}{}_{,0} + \lambda a^{\hat{k}} T^{\hat{0}\hat{0}} + \lambda a_{\hat{i}} T^{\hat{k}\hat{i}} - \lambda a_{\hat{i}} n^{\hat{i}} T^{\hat{k}\hat{0}} \\ &- \lambda a^{\hat{k}} n_{\hat{i}} T^{\hat{i}\hat{0}} - \lambda a^{\hat{i}} R T^{\hat{k}\hat{0}}{}_{,\hat{i}} + \lambda a^{\hat{i}} n_{\hat{j}} R T^{\hat{k}\hat{j}}{}_{,\hat{i}}, \end{aligned} \tag{3.100a}$$

$$\begin{aligned} \lambda F^{(\mathrm{ext})\hat{0}\hat{i}} \tilde{j}_{\hat{i}} =& T^{\hat{i}\hat{0}}{}_{,\hat{i}} + \lambda T^{\hat{0}\hat{0}}{}_{,0} - \lambda \hat{n}_{\hat{i}} T^{\hat{i}\hat{0}}{}_{,0} + 2\lambda a_{\hat{i}} T^{\hat{i}\hat{0}} - \lambda a_{\hat{i}} \hat{n}^{\hat{i}} T^{\hat{0}\hat{0}} \\ &- \lambda a_{\hat{i}} \hat{n}_{\hat{j}} T^{\hat{i}\hat{j}} - \lambda a^{\hat{i}} R T^{\hat{0}\hat{0}}{}_{,\hat{i}} + \lambda a^{\hat{i}} \hat{n}_{\hat{j}} R T^{\hat{j}\hat{0}}{}_{,\hat{i}}, \end{aligned} \tag{3.100b}$$

and

$$0 = \delta^{i}{}_{\hat{j}} j^{\hat{j}}{}_{,i} + \lambda j^{\hat{0}}{}_{,0} - \lambda n^{\hat{i}} \delta_{\hat{i}\hat{j}} j^{\hat{j}}{}_{,0} + \lambda a_{\hat{i}} j^{\hat{i}} - \lambda a_{\hat{i}} j^{\hat{0}} n^{\hat{i}} - \lambda a^{\hat{i}} R\, \delta^{j}{}_{\hat{i}} j^{\hat{0}}{}_{,j} + \lambda a^{\hat{i}} n^{\hat{j}} R\, \delta_{\hat{j}\hat{k}} \delta^{l}{}_{\hat{i}} j^{\hat{k}}{}_{,l}, \tag{3.101}$$

where $f_{,0}$ means $\partial f/\partial\tau$ and $\partial_{\hat{i}} f$ means $\partial f/\partial y^{\hat{i}}$.

We next multiply (3.100) and (3.101) by $R^m n^{\hat{j}_1} \dots n^{\hat{j}_N}$ for integers m and N and integrate with respect to Y. this gives the hierarchy of moment equations

$$\int d^3Y \nabla_\mu \tilde{T}^{\hat{i}\mu} R^m n^{\hat{j}_1} \dots n^{\hat{j}_N} = \int d^3Y F^{(\mathrm{ext})\hat{i}\mu} j_\mu R^m n^{\hat{j}_1} \dots n^{\hat{j}_N}, \tag{3.102a}$$

$$\int d^3Y \nabla_\mu \tilde{T}^{\hat{0}\mu} R^m n^{\hat{j}_1} \dots n^{\hat{j}_N} = \int d^3Y F^{(\mathrm{ext})\hat{0}\mu} j_\mu R^m n^{\hat{j}_1} \dots n^{\hat{j}_N}, \tag{3.102b}$$

$$\int d^3Y \nabla_\mu \tilde{j}^{\mu} R^m n^{\hat{j}_1} \dots n^{\hat{j}_N} = 0. \tag{3.102c}$$

In these equations the arguments of all of the functions are $(\lambda, \tau, Y^{\hat{i}})$, except for $F^{(\mathrm{ext})\hat{a}\hat{b}}$, for which the arguments are as on the right hand side of Eq. (3.51).

We now expand the λ-dependence of $\tilde{T}^{\hat{a}\hat{b}}$ and $\tilde{j}^{\hat{a}}$ at fixed $(\tau, Y^{\hat{i}})$ as

$$\tilde{T}^{\hat{a}\hat{b}} = \tilde{T}^{(0)\hat{a}\hat{b}} + \lambda \tilde{T}^{(1)\hat{a}\hat{b}} + \mathcal{O}(\lambda^2) \tag{3.103a}$$

$$\tilde{j}^{\hat{a}} = \tilde{j}^{(0)\hat{a}} + \lambda \tilde{j}^{(1)\hat{a}} + \mathcal{O}(\lambda^2), \tag{3.103b}$$

with corresponding expansion of the rescaled moments

$$\tilde{P}^{\hat{a}} = \tilde{P}^{(0)\hat{a}} + \lambda \tilde{P}^{(1)\hat{a}} + \mathcal{O}(\lambda^2), \tag{3.104}$$

and similarly for each of the spin (3.91b) and the electromagnetic moments (3.92).

The first moments of the spatial component (3.102a) at leading order, after integrating the spatial partial derivative $\partial_{\hat{i}}$ by parts, and obtaining a boundary term, are

$$-\int d^3Y\, n^{\hat{i}} \tilde{T}^{(0)\hat{k}\hat{j}} \delta_{\hat{i}\hat{j}} = 0 \qquad (m = 1, N = 0), \tag{3.105a}$$

$$-\int d^3Y\, \tilde{T}^{(0)\hat{k}\hat{i}} = 0 \qquad (m = 1, N = 1), \tag{3.105b}$$

$$-\int d^3Y\, R \tilde{T}^{(0)\hat{k}\hat{l}} - \int d^3Y\, n^{\hat{l}} n^{\hat{i}} R \tilde{T}^{(0)\hat{k}\hat{j}} \delta_{\hat{i}\hat{j}}$$
$$-\tfrac{1}{6} \left(\tilde{\mathcal{J}}^{(0)\hat{0}}\right)^2 \delta^{\hat{k}\hat{l}} = 0 \qquad (m = 2, N = 1), \tag{3.105c}$$

$$-\int d^3Y\, n^{\hat{j}} R \tilde{T}^{(0)\hat{k}\hat{i}} - \int d^3Y\, n^{\hat{i}} R \tilde{T}^{(0)\hat{k}\hat{j}} = 0 \qquad (m = 2, N = 2). \tag{3.105d}$$

The boundary terms can be evaluated using Eqs. (3.99),(3.50),(3.9), and (3.53) and are nonzero only in (3.105c).

The first moments of the time component (3.102b) yield

$$-\int d^3Y\, n^{\hat{i}}\tilde{T}^{(0)\hat{j}\hat{0}}\delta_{\hat{i}\hat{j}} = 0 \qquad (m=1, N=0), \tag{3.106a}$$

$$-\int d^3Y\, \tilde{T}^{(0)\hat{i}\hat{0}} = 0 \qquad (m=1, N=1), \tag{3.106b}$$

$$-\int d^3Y\, R\tilde{T}^{(0)\hat{k}\hat{0}} - \int d^3Y\, n^{\hat{k}}n^{\hat{i}}R\tilde{T}^{(0)\hat{j}\hat{0}}\delta_{\hat{i}\hat{j}} = 0 \qquad (m=2, N=1), \tag{3.106c}$$

$$-\int d^3Y\, n^{\hat{j}}R\tilde{T}^{(0)\hat{i}\hat{0}} - \int d^3Y\, n^{\hat{i}}R\tilde{T}^{(0)\hat{j}\hat{0}} = 0 \qquad (m=2, N=2), \tag{3.106d}$$

(3.106e)

It follows from (3.105a), (3.106b), and (3.91a) that

$$\tilde{P}^{\mu} = \tilde{m}u^{\mu} + \mathcal{O}(\lambda). \tag{3.107}$$

The first moments of (3.101) yield

$$-\int d^3Y n^{\hat{i}}\tilde{j}^{\hat{j}}\delta_{\hat{i}\hat{j}} = 0 \qquad (m=1, N=0), \tag{3.108a}$$

$$-\int d^3Y\, j^{(0)\hat{i}} = 0 \qquad (m=1, N=1), \tag{3.108b}$$

$$-\int d^3Y j^{(0)\hat{j}}n^{\hat{i}}R - \int d^3Y j^{(0)\hat{i}}n^{\hat{j}}R = 0 \qquad (m=2, N=2), \tag{3.108c}$$

$$-\int d^3Y j^{(0)\hat{k}}R^2 - 2\int d^3Y j^{(0)\hat{i}}n^{\hat{k}}n^{\hat{j}}R^2\delta_{\hat{i}\hat{j}} = 0 \qquad (m=3, N=1), \tag{3.108d}$$

$$-\int d^3Y j^{(0)\hat{k}}n^{\hat{i}}n^{\hat{j}}R^2 - \int d^3Y j^{(0)\hat{j}}n^{\hat{i}}n^{\hat{k}}R^2 - \int d^3Y j^{(0)\hat{i}}n^{\hat{j}}n^{\hat{k}}R^2 = 0 \qquad (m=3, N=3). \tag{3.108e}$$

It follows from Eqs. (3.108a),(3.108b), and (3.92a) that

$$\tilde{\mathcal{J}}^{\hat{a}} = \tilde{q}u^{\hat{a}} + \mathcal{O}(\lambda). \tag{3.109}$$

This process may be continued to each higher order in λ. At first order in λ, from the $(m=0, N=0)$ piece of (3.102a) we obtain

$$\begin{aligned} 0 =& F^{(\mathrm{ext})\hat{k}\hat{0}}\int d^3Y\, \tilde{j}^{(0)\hat{0}} + a^{(0)\hat{k}}\int d^3Y\, \tilde{T}^{(0)\hat{0}\hat{0}} - F^{(\mathrm{ext})\hat{k}\hat{j}}\int d^3Y\, \tilde{j}^{(0)\hat{i}}\delta_{\hat{i}\hat{j}} \\ & - a^{(0)\hat{k}}\int d^3Y\, n^{\hat{i}}\tilde{T}^{(0)\hat{j}\hat{0}}\delta_{\hat{i}\hat{j}} + \int d^3Y\, \tilde{T}^{(0)\hat{k}\hat{0}}{}_{,0} - \delta_{\hat{i}\hat{j}}\int d^3Y\, n^{\hat{i}}\tilde{T}^{(0)\hat{k}\hat{j}}{}_{,0}, \end{aligned} \tag{3.110}$$

where the external field is evaluated on the worldline. Combining (3.110) with (3.91),(3.92),(3.105a),(3.106c), and (3.108b) gives,

$$\partial_\tau \tilde{P}^{(0)\hat{i}} + \tilde{P}^{(0)\hat{0}} a^{(0)\hat{i}} = D_\tau \tilde{P}^{(0)\hat{i}} = -F^{(\text{ext})\hat{i}\hat{0}} \tilde{\mathcal{J}}^{(0)\hat{0}}. \tag{3.111}$$

Similarly, the $\mathcal{O}(\lambda)$ piece of the $(m = 0, N = 0)$ piece of Eq.(3.102b) together with (3.109) and (3.107) gives

$$\partial_\tau \tilde{m} = \mathcal{O}(\lambda). \tag{3.112}$$

Combining this with (3.111) gives

$$\tilde{m} a^{(0)\hat{i}} = -F^{(\text{ext})\hat{i}\hat{0}} \tilde{\mathcal{J}}^{(0)\hat{0}}, \tag{3.113}$$

the Lorentz force law.

This procedure may be extended to higher moments, and to higher orders in perturbation theory to yield the self force expressions in Secs. 3.6.4-3.6.5, giving the final results presented in Sec. 3.5.2.

The computation of the set of equations (3.102) was automated, using the Mathematica computer algebra software. The notebook used to compute the self force can be found at [193]. The equations we present take advantage of the worldline-based tetrads in the retarded coordinates to re-assemble a covariant form for the laws of motion, so retarded coordinates appear nowhere in our final results in section 3.5. The hierarchy of equations (3.102) is similar to that used by GHW, except that they use integrals over spacelike hypersurfaces

3.6.4 First order laws of motion: Abraham-Lorentz-Dirac

3.6.4.1 Derivation of law of motion

To derive the first order laws of motion, we expand the scaled field equations (3.100) and (3.101) to second order in λ. We will need to use the spin supplementary condition for the first order laws of motion, so we'll present first the leading self-torque, and we will derive the required spin renormalization (3.61) from the leading order self-torque.

We first compute the component of the bare momentum orthogonal to the worldline through $\mathcal{O}(\lambda)$ by combining the $(m = 1, N = 0)$ piece of (3.102a) at $\mathcal{O}(\lambda)$ with the $(m = 1, N = 1)$ piece of

(3.102b), together with (3.91), (3.92). The result is

$$\tilde{P}^{\eta}\mathcal{P}_{\eta}{}^{\mu} = -\lambda\frac{2}{3}\tilde{q}^2 a^{\mu} + \lambda\mathcal{P}^{\mu}{}_{\eta}D_{\tau}\tilde{S}^{\eta\nu}u_{\nu} + 2\lambda\mathcal{P}^{\mu}{}_{\eta}F^{(\mathrm{ext})[\eta}{}_{\lambda}\tilde{Q}^{\lambda|\nu]}u_{\nu} + \mathcal{O}(\lambda^2). \tag{3.114}$$

Here we have converted from equations involving tetrad components to covariant equations, by using the fact that derivatives with respect to τ of tetrad components evaluate on the worldline can be converted to covariant Fermi derivatives $D_F/d\tau$ [159], defined for any vector v^{μ} by

$$\frac{D_F}{d\tau}v^{\mu} = \frac{D}{d\tau}v^{\mu} + (a^{\mu}u^{\nu} - a^{\nu}u^{\mu})v_{\nu}. \tag{3.115}$$

We also note that Eq. (3.114) could equivalently have been derived directly from (3.42) instead of by taking moments of the field equation.

We next compute the first covariant derivative of both the bare momentum and the bare spin through $\mathcal{O}(\lambda^2)$. The covariant derivative of the bare momentum is obtained from the $(m = 0, N = 0)$ moment of the equations (3.102a,3.102b) and the covariant derivative of the spin is obtained from the antisymmetrized moment (3.102a) $(m = 1, N = 1)$.

$$D_{\tau}\tilde{P}^{\lambda} = F^{(\mathrm{ext})\lambda\mu}\tilde{\mathcal{J}}_{\mu} + \lambda F^{(\mathrm{ext})\lambda}{}_{\mu;\nu}\tilde{Q}^{\mu\nu} - \lambda\frac{2}{3}\tilde{q}^2 a_{\nu}a^{\nu}u^{\lambda} + \mathcal{O}(\lambda^2), \tag{3.116a}$$

$$D_{\tau}\tilde{S}^{\mu\nu} = F^{(\mathrm{ext})[\mu}{}_{\lambda}\tilde{Q}^{\lambda\nu]} + \mathcal{O}(\lambda). \tag{3.116b}$$

We also expand the rest mass, which contains no new correction at this order, by combining (3.107), (3.57), and (3.114). The result is

$$\tilde{m} = -\tilde{P}^{\mu}u_{\mu} + \mathcal{O}(\lambda^2). \tag{3.117}$$

At this point, we have imposed no spin supplementary condition, so these equations are entirely general[4], but do not describe the evolution of a worldline. To compute the center of mass acceleration, we use the spin supplementary condition (3.26), which reduces at this order to, from Eq. (3.61)

$$\tilde{S}^{\mu\nu}u_{\nu} = \mathcal{O}(\lambda). \tag{3.118}$$

[4]To this order in perturbation theory, and provided the definitions given in section 3.4.

Combining Eqs. (3.114)-(3.117), we deduce the acceleration and evolution of the rest mass:

$$a^\sigma \tilde{m} = \mathcal{P}^\sigma{}_\mu \Big[F^{(\text{ext})\mu\nu}\tilde{\mathcal{J}}_\nu + \lambda F^{(\text{ext})\mu}{}_{\lambda;\nu}\tilde{Q}^{\lambda\nu} + \lambda\tfrac{2}{3}\tilde{q}^2 D_\tau a^\mu + D_\tau\left(a_\lambda \tilde{S}^{\mu\lambda} + u_\nu F^{(\text{ext})[\nu|}{}_\rho \tilde{Q}^{\rho|\mu]}\right)\Big] + \mathcal{O}(\lambda^2), \tag{3.119a}$$

$$D_\tau \tilde{m} = -u_\mu F^{(\text{ext})\mu\lambda}\tilde{\mathcal{J}}_\lambda + u_\mu F^{(\text{ext})\mu}{}_{\lambda;\eta}\tilde{Q}^{\lambda\eta} - 2a_\eta F^{(\text{ext})[\eta}{}_\lambda \tilde{Q}^{\lambda|\nu]}u_\nu + \mathcal{O}(\lambda^2). \tag{3.119b}$$

In addition, we find from the $(m = 0, N = 0)$ component of the charge conservation equation (3.102c) at $\mathcal{O}(\lambda)$ that

$$D_\tau \tilde{q} = \mathcal{O}(\lambda), \tag{3.120}$$

consistent with the fact that charge is conserved to all orders. From the $(m = 1, N = 0)$ and $(m = 1, N = 1)$ pieces together with Eq. (3.92), we find the expression for the charge moment to this order,

$$\tilde{\mathcal{J}}^\mu = \tilde{q}u^\mu + \lambda u_\nu D_\tau \tilde{Q}^{\mu\nu} - \lambda a^\mu u_\nu u_\lambda \tilde{Q}^{\nu\lambda} - \lambda \mathcal{P}^\mu{}_\nu u_\lambda D_\tau \tilde{Q}^{\lambda\nu} + \mathcal{O}(\lambda^2). \tag{3.121}$$

We next rewrite our results (3.116a),(3.119a), and (3.119b) in terms of the projected, renormalized body parameters (3.59) - (3.63) and eliminate $\tilde{\mathcal{J}}^\mu$ using (3.121). This yields the results (3.72) and the leading piece of (3.79) given in the previous section.

3.6.4.2 Consistency check using the Harte equation of motion

We now preform the consistency check described in Sec. 3.3.2. The radiative self field $F_R^{\mu\nu}$ in Eq. (3.45) is given by [174] and [160], for which the only non-vanishing component is

$$\hat{F}^{\eta\sigma}\mathcal{P}_\eta{}^\lambda u_\sigma = F^{(\text{ext})\eta\sigma}\mathcal{P}_\eta{}^\lambda u_\sigma + \lambda\frac{2}{3}\tilde{q}D_\tau a^\eta \mathcal{P}_\eta{}^\lambda + \mathcal{O}(\lambda^2). \tag{3.122}$$

The self stress energy tensor can also be computed from Eq. (3.99); see also Eq. (120) of GHW. Substituting into Eq. (3.45) gives that,

$$D_\tau P_H^\mu - D_\tau P_B^\mu = \lambda^2 D_\tau\left(\frac{2}{3}\tilde{q}^2 a^\mu\right) + \mathcal{O}(\lambda^3), \tag{3.123}$$

and so the right hand side is indeed a total derivative, as required.

3.6.5 New result: second order laws of motion

3.6.5.1 Derivation of laws of motion

The derivation at second order parallels the derivation given above at first order. We follow the same steps as before, to one higher order in λ. First, we derive the bare momentum orthogonal to the worldline from moments $(m=1,N=0)$ of (3.102a) and $(m=1,N=1)$ of (3.102b) through second order. After simplifying according to equations obtained from the full set of moments from $\mathcal{O}(\lambda^2)$ equations, we obtain

$$\begin{aligned}\tilde{P}^{\kappa}\mathcal{P}_{\kappa}{}^{\mu} = \lambda\mathcal{P}^{\mu}{}_{\kappa}\Big[&-\frac{2}{3}\tilde{q}^2 a^{\kappa} + D_\tau\tilde{S}^{\kappa\nu}u_\nu + 2F^{(\mathrm{ext})[\kappa}{}_{\lambda}\tilde{Q}^{\lambda|\nu]}u_\nu + 4\lambda\tilde{Q}^{\lambda\nu[\kappa}F^{(\mathrm{ext})\sigma]}{}_{\lambda;\nu}u_\sigma \\ &+ 3\lambda\tilde{q}a_\nu D_\tau\tilde{Q}^{\nu\kappa} - \tfrac{1}{3}\lambda\tilde{q}a_\nu D_\tau\tilde{Q}^{\kappa\nu} + \lambda\tilde{q}\tilde{Q}^{\kappa\nu}\left(\tfrac{1}{3}u_\nu a_\sigma a^\sigma - \tfrac{2}{3}D_\tau a_\nu\right) \\ &+ \tfrac{4}{3}\lambda\tilde{q}D_\tau a^{\kappa}\tilde{Q}^{\nu\lambda}u_\nu u_\lambda + \tfrac{1}{5}\lambda\tilde{q}a_\nu a^{\nu}\tilde{Q}^{\lambda\kappa}u_\lambda + \tfrac{4}{3}\lambda\tilde{q}D_\tau^2\tilde{Q}^{\nu\kappa}u_\nu \\ &+ \lambda\tilde{q}a^{\kappa}\left(\tfrac{8}{3}D_\tau\tilde{Q}^{\nu\lambda}u_\nu u_\lambda + \tfrac{7}{5}a_\nu\tilde{Q}^{\lambda\nu}u_\lambda + 3a_\nu\tilde{Q}^{\nu\lambda}u_\lambda\right)\Big] + \mathcal{O}(\lambda^3). \end{aligned} \tag{3.124}$$

The higher-order moments fix also the first covariant derivatives of the bare moments. The first derivative of the bare momentum arises from the $(m=0,N=0)$ moment of the equations (3.102a,3.102b), and subsequent simplifications from $\mathcal{O}(\lambda^2)$ moments, and takes the value

$$\begin{aligned}\mathcal{P}^{\sigma}{}_{\kappa}D_\tau\tilde{P}^{\kappa} =&\mathcal{P}^{\sigma}{}_{\kappa}\Big[F^{\kappa\nu}\tilde{\mathcal{J}}_\mu + \lambda F^{(\mathrm{ext})\kappa}{}_{\lambda;\nu}\tilde{Q}^{\lambda\nu} + \tfrac{1}{2}\lambda^2 F^{(\mathrm{ext})\kappa}{}_{\nu;\lambda\sigma}\tilde{Q}^{\nu\lambda\sigma} \\ &+ \lambda^2\tilde{q}D_\tau a^{\kappa}a_\mu u_\nu\left(\tfrac{1}{5}\tilde{Q}^{\nu\mu} - \tfrac{1}{3}\tilde{Q}^{\mu\nu}\right) + \lambda^2 q a_\mu D_\tau a^{\mu}u_\nu\left(\tfrac{1}{5}\tilde{Q}^{\nu\kappa} + \tfrac{1}{3}\lambda^2\tilde{Q}^{\kappa\nu}\right) \\ &+ \lambda^2\tilde{q}a_\mu a^{\mu}\left(\tfrac{4}{15}D_\tau\tilde{Q}^{\nu\kappa}u_\nu + \tfrac{2}{3}\lambda^2 D_\tau\tilde{Q}^{\kappa\nu}u_\nu + \tfrac{4}{15}a_\nu\tilde{Q}^{\nu\kappa}\right) + \tfrac{2}{3}\lambda^2\tilde{q}a_\mu D_\tau^2\tilde{Q}^{[\kappa\mu]} \\ &+ \lambda^2\tilde{q}a^{\kappa}\big(\tfrac{8}{15}a_\mu D_\tau\tilde{Q}^{\nu\mu}u_\nu - \tfrac{2}{3}a_\mu D_\tau\tilde{Q}^{\mu\nu}u_\nu + \tfrac{4}{5}a_\mu a^{\mu}\tilde{Q}^{\nu\lambda}u_\nu u_\lambda \\ &\qquad - \tfrac{2}{15}D_\tau a_\mu\mathcal{P}^{\mu}{}_{\nu}\tilde{Q}^{\lambda\nu}u_\lambda\big)\Big] + \mathcal{O}(\lambda^3) \end{aligned} \tag{3.125a}$$

$$\begin{aligned}u_\mu D_\tau\tilde{P}^{\mu} =&u_\nu F^{(\mathrm{ext})\nu\mu}\tilde{\mathcal{J}}_\mu + \tfrac{2}{3}\lambda\tilde{q}^2 a_\mu a^{\mu} + \lambda u_\mu F^{(\mathrm{ext})\mu}{}_{\nu;\lambda}\tilde{Q}^{\nu\lambda} + \tfrac{1}{2}\lambda^2 u_\mu F^{(\mathrm{ext})\mu}{}_{\nu;\lambda;\sigma}\tilde{Q}^{\nu\lambda\sigma} \\ &- \lambda^2\tilde{q}a_\mu a^{\mu}\left(\tfrac{8}{3}D_\tau\tilde{Q}^{\nu\lambda}u_\nu u_\lambda + \tfrac{22}{15}a_\nu\tilde{Q}^{\lambda\nu}u_\lambda + \tfrac{8}{3}a_\nu\tilde{Q}^{\nu\lambda}u_\lambda\right) \\ &- \tfrac{4}{3}\lambda^2\tilde{q}a_\mu D_\tau a^{\mu}\tilde{Q}^{\nu\lambda}u_\nu u_\lambda - \tfrac{4}{3}\lambda^2\tilde{q}a_\mu D_\tau^2\tilde{Q}^{\nu\mu}u_\nu \\ &- \tfrac{2}{3}\lambda^2\tilde{q}a_\mu D_\tau a_\nu\tilde{Q}^{\nu\mu} + \mathcal{O}(\lambda^4). \end{aligned} \tag{3.125b}$$

The torque is computed from the antisymmetric part $(m=1,N=1)$ of (3.102a) and simplifications from $\mathcal{O}(\lambda^2)$ equations,

$$\begin{aligned}D_\tau \tilde{S}^{\nu\lambda}\mathcal{P}_\nu{}^\kappa\mathcal{P}_\lambda{}^\mu =&2F^{(\text{ext})[\kappa|}{}_\nu\tilde{Q}^{\nu|\mu]}+2\lambda F^{(\text{ext})[\kappa|}{}_{\nu;\lambda}\tilde{Q}^{\nu\lambda|\mu]}+2\lambda\tilde{q}a^{[\kappa}\mathcal{P}^{\mu]}{}_\lambda\left(\tfrac{1}{3}D_\tau\tilde{Q}^{\nu\lambda}u_\nu+\tfrac{2}{3}D_\tau\tilde{Q}^{\lambda\nu}u_\nu\right)\\&+\tfrac{2}{3}\lambda\tilde{q}D_\tau a^{[\kappa}\tilde{Q}^{\mu]\nu}u_\nu-\tfrac{2}{3}\lambda\tilde{q}D_\tau^2\tilde{Q}^{[\kappa\mu]}+\mathcal{O}(\lambda^2).\end{aligned}\tag{3.126}$$

The rest mass is derived by expanding 3.57, using the bare momentum (3.114). This gives

$$\begin{aligned}\tilde{m}+\tilde{P}^\mu u_\mu=\lambda^2\frac{1}{\tilde{m}}\Bigg(&-\frac{2}{9}\tilde{q}^4a_\mu a^\mu+\frac{8}{3}\tilde{q}^2a_\mu F^{(\text{ext})[\mu|}{}_\nu\tilde{Q}^{\nu|\lambda]}u_\lambda\\&-\frac{1}{2}a_\mu a_\nu\tilde{S}^\mu{}_\kappa\tilde{S}^{\nu\kappa}-a_\nu F^{(\text{ext})}{}_\mu{}^{[\lambda|}\tilde{Q}^{\mu|\eta]}\tilde{S}^\nu{}_\lambda u_\eta\\&-\frac{1}{2}F^{(\text{ext})}{}_\kappa{}^{[\lambda|}F^{(\text{ext})\mu}{}_{[\nu|}\tilde{Q}^{\kappa|\sigma]}\tilde{Q}_{\mu|\eta]}u_\sigma u^\eta\mathcal{P}^\nu{}_\lambda\Bigg)+\mathcal{O}(\lambda^3).\end{aligned}\tag{3.127}$$

Similarly, we derive the charge moment through second order using the $(m=1,N=0)$ and $(m=1,N=1)$ pieces of Eq.(3.102c) at $\mathcal{O}(\lambda^2)$. The result is

$$\begin{aligned}\tilde{\mathcal{J}}^\mu=&\tilde{q}u^\mu+\lambda u_\nu D_\tau\tilde{Q}^{\mu\nu}-\lambda a^\mu u_\nu u_\lambda Q^{\nu\lambda}-\lambda\mathcal{P}^\mu{}_\nu u_\lambda D_\tau\tilde{Q}^{\lambda\nu}-\tfrac{1}{2}\lambda^2D_\tau a^\mu\tilde{Q}^{\nu\lambda\rho}u_\nu u_\lambda u_\rho\\&-\tfrac{1}{2}\lambda^2D_\tau^2\tilde{Q}^{\mu\nu\lambda}u_\nu u_\lambda-\lambda^2a^\mu\left(\tfrac{3}{2}a_\nu\tilde{Q}^{(\nu\lambda\rho)}u_\lambda u_\rho+\tfrac{3}{2}D_\tau\tilde{Q}^{\nu\lambda\rho}u_\nu u_\lambda u_\rho\right)\\&-\lambda^2\mathcal{P}^\mu{}_\nu\left(3D_\tau\tilde{Q}^{(\nu\lambda\rho)}a_\lambda u_\rho+D_\tau^2\tilde{Q}^{\lambda\nu\rho}u_\lambda u_\rho\right)+\mathcal{O}(\lambda^3).\end{aligned}\tag{3.128}$$

Finally, to evaluate the explicit equations of motion for the worldline and for the evolution of the rest mass, we use the following rescaled versions of the general identities (3.38):

$$\tilde{m}a^\kappa=a^\kappa\left(\tilde{m}+\tilde{P}^\mu u_\mu\right)+\mathcal{P}^\kappa{}_\lambda D_\tau\tilde{P}^\lambda-\mathcal{P}^\kappa{}_\nu D_\tau\left(\mathcal{P}^\nu{}_\lambda\tilde{P}^\lambda\right),\tag{3.129a}$$

$$D_\tau\tilde{m}=D_\tau\left(\tilde{m}+\tilde{P}^\mu u_\mu\right)-u_\mu D_\tau\tilde{P}^\mu-a_\mu\tilde{P}^\mu,\tag{3.129b}$$

One can think of the first and third terms in each of (3.129) as representing the effect of hidden momentum, that is, the component of momentum perpendicular to $\vec{u}$. By substituting the results (3.124)-(3.127) and (3.128) into the general identity (3.129), making use of the spin supplementary condition (3.26), and eliminating the body parameters in terms of the renormalized projected body parameters (3.59)-(3.66), we finally arrive at the second order equations of motion (3.75)-(3.78).

3.6.5.2 Consistency check using the Harte equation of motion

We turn now to the consistency check described in Sec. 3.3.2. We first compute the regular self field through second order. The expansions (3.99) we use to derive these expressions are expanded

asymptotically at large R. However, taking the difference between the retarded and advanced fields in the multipole expansion, re-expanded at small R will yield the regular field. This procedure can be thought of as obtaining an asymptotic form for the fields, then replacing the extended source with a pointlike source for the purposes of computing the regular field. Since the regular field should depend only on the standard multipoles of the body, the regular field should be indistinguishable for the extended and replaced pointlike body. This procedure and argument are analogous to that used by Pound [194] for the gravitational case.

The result in terms of the tetrad components and retarded coordinates from 3.6.1 is

$$\begin{aligned}
F_{R\,\hat{k}\hat{0}} =& \tfrac{2}{3}\lambda D_\tau a^{\hat{i}}\tilde{q}\delta_{\hat{k}\hat{i}} - \tfrac{2}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}a_{\hat{k}}\tilde{q}R - \tfrac{2}{3}\lambda^2 a_{\hat{k}}D_\tau a^{\hat{i}}\tilde{q}n^{\hat{j}}R\delta_{\hat{i}\hat{j}} + \tfrac{2}{3}\lambda^2 {D_\tau}^2 a^{\hat{i}}\tilde{q}R\delta_{\hat{k}\hat{i}} \\
&- \tfrac{4}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{j}}\tilde{q}n^{\hat{i}}R\delta_{\hat{k}\hat{j}} + \tfrac{2}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{i}}\tilde{q}n^{\hat{j}}R\delta_{\hat{k}\hat{j}} - \tfrac{2}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{i}}\tilde{Q}^{\hat{0}\hat{j}}\delta_{\hat{k}\hat{j}} - \tfrac{2}{3}\lambda^2 {D_\tau}^2 a^{\hat{i}}\tilde{Q}^{\hat{j}\hat{l}}\delta_{\hat{i}\hat{l}}\delta_{\hat{k}\hat{j}} \\
&+ \tfrac{1}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}a_{\hat{j}}\tilde{Q}^{\hat{j}\hat{l}}\delta_{\hat{k}\hat{l}} + \tfrac{2}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}a_{\hat{j}}\tilde{Q}^{\hat{l}\hat{j}}\delta_{\hat{k}\hat{l}} - \tfrac{1}{3}\lambda^2 {D_\tau}^2 a^{\hat{i}}\tilde{Q}^{\hat{j}\hat{l}}\delta_{\hat{i}\hat{j}}\delta_{\hat{k}\hat{l}} - \tfrac{2}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}\delta_{\hat{k}\hat{j}}\partial_\tau\tilde{Q}^{\hat{0}\hat{j}} \\
&- \tfrac{5}{3}\lambda^2 D_\tau a^{\hat{i}}\delta_{\hat{i}\hat{l}}\delta_{\hat{k}\hat{j}}\partial_\tau\tilde{Q}^{\hat{j}\hat{l}} - \lambda^2 D_\tau a^{\hat{i}}\delta_{\hat{i}\hat{l}}\delta_{\hat{k}\hat{j}}\partial_\tau\tilde{Q}^{\hat{l}\hat{j}} - \lambda^2 a_{\hat{i}}\delta_{\hat{k}\hat{j}}{\partial_\tau}^2\tilde{Q}^{\hat{i}\hat{j}} - \lambda^2 a_{\hat{i}}\delta_{\hat{k}\hat{j}}{\partial_\tau}^2\tilde{Q}^{\hat{j}\hat{i}} \\
&+ \tfrac{2}{3}\lambda^2\delta_{\hat{k}\hat{i}}{\partial_\tau}^3\tilde{Q}^{\hat{0}\hat{i}} + \mathcal{O}(\lambda^3),
\end{aligned} \tag{3.130}$$

and

$$\begin{aligned}
F_{R\,\hat{k}\hat{j}} =& \tfrac{2}{3}\lambda^2 a_{\hat{k}}D_\tau a^{\hat{i}}\tilde{q}R\delta_{\hat{j}\hat{i}} - \tfrac{1}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}a_{\hat{k}}\tilde{q}n^{\hat{l}}R\delta_{\hat{j}\hat{l}} + \tfrac{1}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}a_{\hat{k}}\tilde{Q}^{\hat{0}\hat{l}}\delta_{\hat{j}\hat{l}} - \tfrac{2}{3}\lambda^2 a_{\hat{k}}D_\tau a^{\hat{i}}\tilde{Q}^{\hat{l}\hat{m}}\delta_{\hat{i}\hat{m}}\delta_{\hat{j}\hat{l}} \\
&- \tfrac{2}{3}\lambda^2 a_{\hat{k}}D_\tau a^{\hat{i}}\tilde{Q}^{\hat{l}\hat{m}}\delta_{\hat{i}\hat{l}}\delta_{\hat{j}\hat{m}} - \tfrac{2}{3}\lambda^2 a_{\hat{j}}D_\tau a^{\hat{i}}\tilde{q}R\delta_{\hat{k}\hat{i}} + \tfrac{1}{3}\lambda^2 {D_\tau}^2 a^{\hat{i}}\tilde{q}n^{\hat{l}}R\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{i}} - \tfrac{1}{3}\lambda^2 {D_\tau}^2 a^{\hat{i}}\tilde{Q}^{\hat{0}\hat{l}}\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{i}} \\
&+ \tfrac{1}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}a_{\hat{j}}\tilde{q}n^{\hat{l}}R\delta_{\hat{k}\hat{l}} - \tfrac{1}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}a_{\hat{j}}\tilde{Q}^{\hat{0}\hat{l}}\delta_{\hat{k}\hat{l}} + \tfrac{2}{3}\lambda^2 a_{\hat{j}}D_\tau a^{\hat{i}}\tilde{Q}^{\hat{l}\hat{m}}\delta_{\hat{i}\hat{m}}\delta_{\hat{k}\hat{l}} - \tfrac{1}{3}\lambda^2 {D_\tau}^2 a^{\hat{i}}\tilde{q}n^{\hat{l}}R\delta_{\hat{j}\hat{i}}\delta_{\hat{k}\hat{l}} \\
&+ \tfrac{1}{3}\lambda^2 {D_\tau}^2 a^{\hat{i}}\tilde{Q}^{\hat{0}\hat{l}}\delta_{\hat{j}\hat{i}}\delta_{\hat{k}\hat{l}} - \tfrac{1}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{l}}\tilde{Q}^{\hat{i}\hat{m}}\delta_{\hat{j}\hat{m}}\delta_{\hat{k}\hat{l}} + \tfrac{1}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{i}}\tilde{Q}^{\hat{l}\hat{m}}\delta_{\hat{j}\hat{m}}\delta_{\hat{k}\hat{l}} \\
&- \tfrac{1}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{l}}\tilde{Q}^{\hat{m}\hat{i}}\delta_{\hat{j}\hat{m}}\delta_{\hat{k}\hat{l}} + \tfrac{2}{3}\lambda^2 a_{\hat{j}}D_\tau a^{\hat{i}}\tilde{Q}^{\hat{l}\hat{m}}\delta_{\hat{i}\hat{l}}\delta_{\hat{k}\hat{m}} + \tfrac{1}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{l}}\tilde{Q}^{\hat{i}\hat{m}}\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{m}} \\
&- \tfrac{1}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{i}}\tilde{Q}^{\hat{l}\hat{m}}\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{m}} + \tfrac{1}{3}\lambda^2 a_{\hat{i}}D_\tau a^{\hat{l}}\tilde{Q}^{\hat{m}\hat{i}}\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{m}} - \tfrac{2}{3}\lambda^2 D_\tau a^{\hat{i}}\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{i}}\partial_\tau\tilde{Q}^{\hat{0}\hat{l}} \\
&+ \tfrac{2}{3}\lambda^2 D_\tau a^{\hat{i}}\delta_{\hat{j}\hat{i}}\delta_{\hat{k}\hat{l}}\partial_\tau\tilde{Q}^{\hat{0}\hat{l}} - \lambda^2 a_{\hat{i}}a_{\hat{k}}\delta_{\hat{j}\hat{l}}\partial_\tau\tilde{Q}^{\hat{i}\hat{l}} + \lambda^2 a_{\hat{i}}a_{\hat{j}}\delta_{\hat{k}\hat{l}}\partial_\tau\tilde{Q}^{\hat{i}\hat{l}} - \lambda^2 a_{\hat{i}}a_{\hat{k}}\delta_{\hat{j}\hat{l}}\partial_\tau\tilde{Q}^{\hat{l}\hat{i}} \\
&+ \lambda^2 a_{\hat{i}}a_{\hat{j}}\delta_{\hat{k}\hat{l}}\partial_\tau\tilde{Q}^{\hat{l}\hat{i}} - \tfrac{1}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{m}}\partial_\tau\tilde{Q}^{\hat{l}\hat{m}} + \tfrac{1}{3}\lambda^2 a_{\hat{i}}a^{\hat{i}}\delta_{\hat{j}\hat{l}}\delta_{\hat{k}\hat{m}}\partial_\tau\tilde{Q}^{\hat{m}\hat{l}} + \tfrac{1}{3}\lambda^2\delta_{\hat{j}\hat{i}}\delta_{\hat{k}\hat{l}}{\partial_\tau}^3\tilde{Q}^{\hat{i}\hat{l}} \\
&- \tfrac{1}{3}\lambda^2\delta_{\hat{j}\hat{i}}\delta_{\hat{k}\hat{l}}{\partial_\tau}^3\tilde{Q}^{\hat{l}\hat{i}} + \mathcal{O}(\lambda^3)
\end{aligned} \tag{3.131}$$

Inserting covariant versions of these expressions into the first term on the RHS of Eq. (3.45) gives

$$\begin{aligned}\mathcal{P}^{\kappa}{}_{\nu}\int d^3\Sigma_{\tilde{\mu}} m^{\tilde{\mu}} g^{\nu}{}_{\tilde{\lambda}} F_R^{\tilde{\lambda}\tilde{\rho}} j_{\tilde{\rho}} = \mathcal{P}^{\kappa}{}_{\nu}\Big[&\tfrac{2}{3}\lambda^2\tilde{q}^2 D_\tau a^\nu + \tfrac{2}{3}\lambda^3\tilde{q}a^\nu D_\tau a^\mu \mathcal{P}_{\mu\lambda}\tilde{Q}^{\rho\lambda}u_\rho - \tfrac{2}{3}\lambda^3\tilde{q}a^\nu D_\tau a_\mu \tilde{Q}^{\mu\rho}u_\rho \\ &+ \tfrac{2}{3}\lambda^3\tilde{q}D_\tau{}^2 a^\nu \tilde{Q}^{\lambda\rho}u_\lambda u_\rho + \tfrac{2}{3}\lambda^3\tilde{q}D_\tau a^\nu a_\mu \tilde{Q}^{\rho\mu}u_\rho \\ &+ \tfrac{2}{3}\lambda^3\tilde{q}D_\tau a^\nu D_\tau\left(\tilde{Q}^{\lambda\rho}u_\lambda u_\rho\right) + \tfrac{2}{3}\lambda^3\tilde{q}a_\mu a^\mu \delta_{\hat{k}\hat{j}} D_\tau\left(u_\lambda \mathcal{P}^\nu{}_\rho \tilde{Q}^{\lambda\rho}\right) \\ &- \tfrac{2}{3}\lambda^3\tilde{q}D_\tau a_\sigma \mathcal{P}^\sigma{}_\mu D_\tau\left(\mathcal{P}^\nu{}_\lambda \mathcal{P}^\mu{}_\rho \tilde{Q}^{\lambda\rho}\right) \\ &+ \tfrac{2}{3}\lambda^3\tilde{q}D_\tau\left(\mathcal{P}^\nu{}_\mu D_\tau\left(\mathcal{P}^\mu{}_\lambda D_\tau\left(\mathcal{P}^\lambda{}_\rho u_\sigma \tilde{Q}^{\sigma\rho}\right)\right)\right)\Big] + \mathcal{O}(\lambda^4),\end{aligned} \tag{3.132a}$$

$$\begin{aligned}u_\nu\int d^3\Sigma_{\tilde{\mu}} m^{\tilde{\mu}} g^{\nu}{}_{\tilde{\lambda}} F_R^{\tilde{\lambda}\tilde{\rho}} j_{\tilde{\rho}} = &-\tfrac{2}{3}\lambda^3\tilde{q}a_\mu D_\tau a^\mu \tilde{Q}^{\nu\lambda}u_\nu u_\lambda + \tfrac{2}{3}\lambda^3\tilde{q}D_\tau{}^2 a_\mu \mathcal{P}^\mu{}_\nu \tilde{Q}^{\nu\lambda}u_\lambda \\ &- \tfrac{2}{3}\lambda^3\tilde{q}a_\nu D_\tau a_\mu \tilde{Q}^{\nu\mu} + \tfrac{4}{3}\lambda^3\tilde{q}D_\tau a^\mu \mathcal{P}_{\mu\nu} D_\tau\left(\mathcal{P}^\nu{}_\lambda u_\rho \tilde{Q}^{(\rho\lambda)}\right) \\ &+ \mathcal{O}(\lambda^4)\end{aligned} \tag{3.132b}$$

To evaluate the second term on the right hand side of (3.45), we note from Eqs. (3.22a), (3.34),(3.36), and (3.56a) that it is given by the right hand sides of Eq.(3.125), multiplied by λ, and with the external fields set to zero. Equation (3.45) thus evaluates to

$$\begin{aligned}D_\tau P_H^\kappa - D_\tau P_B^\kappa = \lambda D_\tau\Big[&\tfrac{2}{3}\lambda\tilde{q}^2 a^\kappa - \tfrac{4}{3}\lambda^2\tilde{q}a^\kappa\left(D_\tau\tilde{Q}^{\lambda\eta}u_\lambda u_\eta + \tfrac{2}{5}a_\nu\tilde{Q}^{\lambda\nu}u_\lambda + a_\nu\tilde{Q}^{\nu\lambda}u_\lambda\right) \\ &- \tfrac{8}{3}\lambda^2\tilde{q}a_\nu\mathcal{P}^\kappa{}_\mu D_\tau\tilde{Q}^{[\nu\mu]} + \lambda^2\tilde{q}a_\nu D_\tau\tilde{Q}^{\kappa\nu} - \tfrac{2}{3}\lambda^2 u^\kappa a_\mu a_\nu\tilde{Q}^{\mu\nu} \\ &- \tilde{q}\lambda^2\mathcal{P}^\kappa{}_\mu\left(\tfrac{2}{3}D_\tau^2\tilde{Q}^{\lambda\mu}u_\lambda + \tfrac{4}{15}a_\nu a^\nu\tilde{Q}^{\lambda\kappa}u_\lambda + \tfrac{2}{3}a_\nu a^\nu\tilde{Q}^{\kappa\lambda}u_\lambda\right) \\ &- \tfrac{2}{3}\lambda^2(\mathcal{P}^\kappa{}_\mu + u^\kappa u_\mu)D_\tau a_\nu\tilde{Q}^{\nu\mu}\Big] + \mathcal{O}(\lambda^4).\end{aligned} \tag{3.133}$$

The right hand side is a total derivative as required, so our results satisfy the consistency condition.

3.7 Conclusions

In this paper, we have demonstrated the use of rigorous, limit based methods for deriving higher-order self forces. Via an extension to the method first introduced by GHW, combined with reasoning motivated by the work of Harte [174], we have derived the entire self force effect through second order without any ad hoc regularization. These methods also yield the full multipole dependence

of radiation-reaction effects. The dipole dependence of the first order radiation-reaction force was derived by GHW, and we find the analogous second order dependence on dipole and quadrupole contributions. Our results contain the first extended body dependence of any second order self force, electromagnetic or otherwise, as well as the first explicit expression for the self torque, which first arises at second order.

Acknowledgments

We thank Abe Harte and Justin Vines for helpful conversations. This research was supported in part by NSF grants PHY-1404105 and PHY-1707800.

Appendix

3.A Convergence of integrals for bare spin and momentum

In this appendix, we show that the integral (3.34)

$$P_\tau(\xi) = \int_{\Sigma_\tau} T^{\tilde{\mu}\tilde{\nu}} \xi_{\tilde{\mu}} d\Sigma_{\tilde{\nu}} \tag{3.134}$$

is well defined in Minkowski spacetime when $\xi^{\tilde{\mu}}$ is one of the ten Killing vector fields, Σ_τ is a future null cone, and $T^{\tilde{\mu}\tilde{\nu}}$ is the stress-energy tensor (3.15) that involves the retarded self-field. Different choices of Killing vector field $\xi^{\tilde{\mu}}$ give rise to our definitions (3.22) of linear momentum and spin.

We fix a point z_τ on the center of mass worldline and introduce coordinates $(u, r, \theta, \phi) = (u, r, \theta^1, \theta^2) = (u, r, \theta^A)$ such that the metric is

$$ds^2 = -2dudr - du^2 + r^2 d\Omega^2 \tag{3.135}$$

and that the null cone Σ_τ is the surface $u = \tau =$ constant. We define $n_\mu = -(du)_\alpha$, the null normal to Σ_τ. The integral (3.134) can be written as

$$P_\tau(\xi) \propto \int_0^\infty dr r^2 \int d^2\Omega Q_\mu \xi^\mu, \tag{3.136}$$

where we have dropped the tildes for simplicity and

$$Q_\mu = T_{\mu\nu} n^\nu. \tag{3.137}$$

A priori, we would not expect the integral (3.136) to converge, since the leading order components of $T_{\mu\nu}$ scale as $1/r^2$. However, we shall see that cancellations occur because the surface Σ_τ is asymptotically a surface of constant phase for the outgoing radiation. From Eq.(3.136), a sufficient condition for convergence is that

$$\int d^2\Omega Q_\mu \xi^\mu = \mathcal{O}(r^{-4}) \tag{3.138}$$

as $r \to \infty$.

The general form of a Killing vector field in the coordinates (3.135) as $r \to \infty$ is [195]

$$\vec{\xi} = \left[\alpha + \tfrac{1}{2}u\Psi + \mathcal{O}(r^{-1})\right]\partial_u + \left[Y^A + \mathcal{O}(r^{-1})\right]\partial_A - \left[\tfrac{1}{2}r\Psi + \mathcal{O}(1)\right]\partial_r, \tag{3.139}$$

where $Y^A(\theta^B)$ is a conformal Killing vector field on the 2-sphere that encodes rotation and boosts, $\Psi = D_A Y^A$, and D_A is the covariant derivative operator with respect to the 2-sphere metric h_{AB} defined by $d\Omega^2 = h_{AB}d\theta^A d\theta^B$. The function $\alpha(\theta^B)$ is a linear combination of $l = 0$ and $l = 1$ spherical harmonics and encodes translations.

Now inserting (3.139) into (3.138), we find the sufficient condition for convergence is

$$\int d^2\Omega \Big\{ \left[\tfrac{1}{2}u\Psi + \alpha + \mathcal{O}(r^{-1})\right] Q_u + \left[Y^A + \mathcal{O}(r^{-1})\right] Q_A + \left[-\tfrac{1}{2}r\Psi + \mathcal{O}(1)\right] Q_r \Big\} = \mathcal{O}(r^{-4}), \tag{3.140}$$

which will be satisfied if

$$Q_u = \mathcal{O}(r^{-4}), \tag{3.141a}$$

$$Q_A = \mathcal{O}(r^{-4}), \tag{3.141b}$$

$$Q_r = \mathcal{O}(r^{-5}). \tag{3.141c}$$

Consider first the scalar case. When the scalar charge density ρ is smooth, the method of Sec. 11.1 of [196] can be used to show that the retarded scalar field $\Phi^{(\text{self})}$ has an expansion near future null infinity of the form

$$\Phi^{(\text{self})} = \frac{f(u,\theta^A)}{r} + \frac{g(u,\theta^A)}{r^2} + \mathcal{O}(r^{-3}), \tag{3.142}$$

for some smooth functions f and g. Inserting this expansion into Eqs.(3.10),(3.14),(3.15), and

(3.137) yields

$$Q_r = \frac{-f^2}{r^4} + \mathcal{O}(r^{-5}) \tag{3.143a}$$

$$Q_u = \frac{-1}{2r^4}\left[f^2 + h^{AB} D_A f D_B f\right] + \mathcal{O}(r^{-4}) \tag{3.143b}$$

$$Q_A = \frac{1}{r^3} f D_A f + \mathcal{O}(r^{-4}) \tag{3.143c}$$

It can be seen that these expressions do not satisfy the scalings (3.141). However, inserting the expressions (3.143) into (3.140) and integrating by parts on the two-sphere, we find that the leading order terms cancel and so the condition (3.140) is satisfied.

Turn now to the electromagnetic case. We can use the method of Sec 11.1 of [196] to deduce the asymptotic scaling of the component of the retarded field $F^{(\text{self})}_{\mu\nu}$. Defining $\rho = r^{-1}$, the metric can be written as $ds^2 = \rho^{-2} d\tilde{s}^2$ with

$$d\tilde{s}^2 = -\rho^2 du^2 - 2 du d\rho + d\Omega^2. \tag{3.144}$$

Since the field equations (3.7) are conformally invariant away from sources, $F^{(\text{self})}_{\mu\nu}$ is a solution of the equations in the metric (3.144) and hence is a smooth function of (ρ, u, θ^A) at $\rho = 0$, i.e. on future null infinity. It follows that for general solutions with smooth sources

$$F^{(\text{self})}_{ur} = \mathcal{O}(r^{-2}), \tag{3.145a}$$

$$F^{(\text{self})}_{uA} = \mathcal{O}(1), \tag{3.145b}$$

$$F^{(\text{self})}_{rA} = \mathcal{O}(r^{-2}), \tag{3.145c}$$

$$F^{(\text{self})}_{AB} = \mathcal{O}(1) \tag{3.145d}$$

as $r \to \infty$. From Eqs. (3.9),(3.14),(3.15), and (3.137) we find that

$$Q_r = -\frac{1}{r^2} F^{(\text{self})}_{rA} F^{(\text{self})}_{rB} h^{AB}, \tag{3.146a}$$

$$Q_A = -F^{(\text{self})}_{Ar} F^{\text{self}}_{ur} - \frac{1}{r^2} F^{(\text{self})}_{AB} F^{(\text{self})}_{rC} h^{BC}, \tag{3.146b}$$

$$Q_u = -\frac{1}{2} F^{(\text{self})2}_{ur} - \frac{1}{2r^2} F^{(\text{self})}_{rA} F^{(\text{self})}_{rB} h^{AB} - \frac{1}{r^4} F^{(\text{self})}_{AB} F^{(\text{self})}_{CD} h^{AC} h^{BD}. \tag{3.146c}$$

Inserting the scalings (3.145) into the expressions (3.146) we find that the conditions for convergence (3.141) are satisfied.

3.B Scalar laws of motion

3.B.1 Renormalized scalar moments

As for the electromagnetic case, we find it useful to introduce a renormalized set of moments to describe the scalar charge distribution, modifying the rescaled moments $\tilde{q_S}$,$\tilde{Q}^\mu_S$, and $\tilde{Q}_S{}^{\mu\nu}$ given in Eq. (3.29). Unlike the electromagnetic case, the scalar charge and so may be renormalized [5], so possesses an ambiguity in the chargelike degrees of freedom. The renormalized charge is

$$q_S = \tilde{q}_S + \lambda D_\tau \tilde{Q}_S{}^\mu u_\mu - \lambda^2 D_\tau \left(u_\mu \tilde{Q}_S{}^{\mu\nu} a_\nu \right) + \mathcal{O}(\lambda^3). \tag{3.147}$$

The renormalized projected dipole is

$$Q_S{}^\mu = \mathcal{P}^\mu{}_\nu \left(\tilde{Q}_S{}^\nu + \lambda D_\tau \tilde{Q}_S^{\nu\lambda} u_\lambda \right) + \mathcal{O}(\lambda^2), \tag{3.148}$$

which is explicitly orthogonal to the 4-velocity. We define the renormalized projected quadrupole as

$$Q_S{}^{\mu\nu} = \mathcal{P}^\mu{}_\lambda \mathcal{P}^\nu{}_\sigma \left(\tilde{Q}_S{}^{\lambda\sigma} \right) + \mathcal{O}(\lambda), \tag{3.149}$$

which is explicitly orthogonal to u^μ in both of its indices, $u_\mu Q_S^{\mu\nu} = u_\nu Q_S^{\mu\nu} = 0$.

In addition, as in the electromagnetic case, we find it useful to define a renormalized mass and a renormalized spin. The definitions are

$$\begin{aligned} m + u_\mu \tilde{P}^\mu =& - \lambda u_\nu \Phi^{(\text{ext});\nu} \tilde{Q}^\mu_S u_\mu + \lambda \tilde{q} D_\tau \tilde{q} - \lambda^2 u_\lambda \Phi^{(\text{ext});\lambda}{}_\mu \mathcal{P}^\mu{}_\nu \tilde{Q}_S^{\nu\rho} u_\rho \\ &+ \lambda^2 u_\mu \Phi^{(\text{ext});\mu} a_\nu \tilde{Q}_S^{\nu\lambda} u_\lambda + \tfrac{1}{3} \lambda^2 \tilde{q} a_\mu a^\mu \tilde{Q}^\nu_S u_\nu + \tfrac{1}{3} \lambda^2 \tilde{q} a_\nu D_\tau \tilde{Q}^\nu_S \\ &+ \lambda^2 \tilde{q} D_\tau{}^2 \tilde{Q}^\mu_S u_\mu - \tfrac{2}{3} \lambda^2 a_\mu \tilde{Q}^\mu_S D_\tau \tilde{q} + \lambda^2 D_\tau \left(\tilde{Q}^\mu_S u_\mu \right) D_\tau \tilde{q} + \mathcal{O}(\lambda^3) \end{aligned} \tag{3.150a}$$

$$S^{\mu\nu} = \tilde{S}^{\mu\nu} + 2\lambda \Phi^{(\text{ext});[\mu} \tilde{Q}_S^{\nu]\lambda} u_\lambda + \tfrac{2}{3} \lambda \tilde{q} a^{[\mu} \tilde{Q}_S^{\nu]} + \tfrac{2}{3} \lambda u^{[\nu} D_\tau \left(\tilde{q} \tilde{Q}_S^{\mu]} \right) + \mathcal{O}(\lambda^2). \tag{3.150b}$$

[5] That is, the definition of the charge depends on the choice of hypersurface, so it is natural to allow a redefinition of the charge in order to simplify the equations of motion.

3.B.2 | Scalar self force in terms of renormalized moments

As in the electromagnetic presentation, we decompose the self force and rest mass evolution as

$$ma^\mu = f_S^{(0)\mu} + \lambda f_S^{(1)\mu} + \lambda^2 f_S^{(2)\mu} + \mathcal{O}(\lambda^3) \tag{3.151a}$$

$$D_\tau m = \mathcal{F}_S^{(0)} + \lambda \mathcal{F}_S^{(1)} + \lambda^2 \mathcal{F}_S^{(2)} + \mathcal{O}(\lambda^3) \tag{3.151b}$$

Following similar steps to the electromagnetic derivation, we find the leading force and mass evolution

$$f_S^{(0)\mu} = q_S \mathcal{P}^\sigma{}_\kappa \Phi^{(\mathrm{ext})\kappa} \tag{3.152a}$$

$$\mathcal{F}_S^{(0)} = -\, q_S \Phi^{(\mathrm{ext})\mu} u_\mu, \tag{3.152b}$$

where $\Phi^{(\mathrm{ext})\mu} \equiv \nabla^\mu \Phi^{(\mathrm{ext})}$. The GHW-order scalar self force and mass evolution,

$$\begin{aligned} f_S^{(1)\mu} = \mathcal{P}^\sigma{}_\kappa \Big[& q_S \Phi^{(\mathrm{ext})\kappa} + \Phi^{(\mathrm{ext})\kappa}{}_{;\mu} Q_S{}^\mu + \tfrac{1}{3} D_\tau a^\kappa q_S^2 + a^\kappa q_S D_\tau q_S \\ & - 2 D_\tau \left(Q_S{}^{[\kappa} \Phi^{(\mathrm{ext})\mu]} u_\mu \right) + D_\tau \left(a_\mu S^{\kappa\mu} \right) \Big] \end{aligned} \tag{3.153a}$$

$$\mathcal{F}^{(1)} = -\, q_S \Phi^{(\mathrm{ext})\mu} u_\mu - u_\nu \Phi^{(\mathrm{ext})\nu}{}_{;\mu} Q_S{}^\mu - 2 \Phi^{(\mathrm{ext})\mu} u_\mu a_\nu Q_S^\nu + q_S D_\tau{}^2 q_S \tag{3.153b}$$

These results are new except for the monopole terms, which can be found in [197]. The second-order results can be expressed as a sum of as a sum of dipole and quadrupole contributions:

$$\begin{aligned} f_S^{(2)\mu} &= f_S^{(2)\mu}{}_{\mathrm{dipole}} + f_S^{(2)\mu}{}_{\mathrm{quadrupole}}, \\ \mathcal{F}_S^{(2)} &= \mathcal{F}_S^{(2)}{}_{\mathrm{dipole}} + \mathcal{F}_S^{(2)}{}_{\mathrm{quadrupole}}, \end{aligned} \tag{3.154a}$$

As for the electromagnetic case, there are no explicit monopole terms at this order. The explicit,

new, dipole and quadrupole contributions to the self force are:

$$\begin{aligned} f_S^{(2)\mu}{}_{\text{dipole}} = \mathcal{P}^\sigma{}_\kappa \Big[& -\tfrac{1}{3} q_S a^\kappa D_\tau Q_S{}^\nu a_\nu - \tfrac{1}{3} a_\nu a^\nu \left(q_S D_\tau Q_S{}^\kappa - D_\tau q_S Q_S{}^\kappa \right) \\ & - \tfrac{2}{3} q_S D_\tau a^\kappa a_\mu Q_S{}^\mu - \tfrac{1}{3} Q_S{}^\kappa D_\tau{}^3 q_S + \tfrac{1}{3} q_S D_\tau{}^3 Q_S{}^\kappa \\ & - q_S a^\kappa D_\tau a_\mu Q_S{}^\mu - D_\tau \left(D_\tau q_S D_\tau Q_S{}^\kappa \right) \Big], \end{aligned} \tag{3.155a}$$

$$\begin{aligned} f_S^{(2)\mu}{}_{\text{quadrupole}} = \mathcal{P}^\sigma{}_\kappa \Big[& \tfrac{1}{2} \nabla \Phi^{(\text{ext})\kappa}{}_{;\mu\nu} Q_S{}^{\mu\nu} + \tfrac{1}{2} Q_S{}^\rho{}_\rho D_\tau{}^2 \Phi^{(\text{ext})\kappa} - D_\tau \left(u_\mu \Phi^{(\text{ext})\mu}{}_{;\nu} Q_S{}^{\kappa\nu} \right) \\ & + \Phi^{(\text{ext})\kappa}{}_{;\mu} Q_S{}^{\mu\nu} a_\nu - D_\tau \left(\Phi^{(\text{ext})\mu} u_\mu Q_S{}^{\kappa\nu} a_\nu \right) + \tfrac{1}{2} \Phi^{(\text{ext})\kappa}{}_{;\mu} a^\mu Q_S{}^\rho{}_\rho \\ & - D_\tau \Phi^{(\text{ext})\mu} u_\mu Q_S{}^{\kappa\nu} a_\nu + D_\tau \left(D_\tau Q_S{}^\rho{}_\rho \Phi^{(\text{ext})\kappa} \right) \\ & - 2 \Phi^{(\text{ext})\kappa} a_\mu a_\nu Q_S{}^{\mu\nu} + a^\kappa \Phi^{(\text{ext})\mu} u_\mu D_\tau Q_S{}^\rho{}_\rho \Big], \end{aligned} \tag{3.155b}$$

and the explicit, new, dipole and quadrupole contributions to the mass evolution are

$$\mathcal{F}_S{}^{(2)}_{\text{dipole}} = \tfrac{1}{3} q_S D_\tau a_\mu \mathcal{P}^\mu{}_\nu D_\tau Q_S{}^\nu - a_\mu D_\tau \left(D_\tau q_S Q_S{}^\mu \right) - \tfrac{4}{3} D_\tau q_S D_\tau a_\mu Q_S{}^\mu, \tag{3.156a}$$

$$\begin{aligned} \mathcal{F}_S{}^{(2)}_{\text{quadrupole}} = & - \Phi^{(\text{ext})\mu} u_\mu a_\nu a_\lambda Q_S{}^{\nu\lambda} - 2 u_\lambda \Phi^{(\text{ext})\lambda}{}_{;\mu} Q_S{}^{\mu\nu} a_\nu - \tfrac{1}{2} u_\mu \Phi^{(\text{ext})\mu}{}_{;\nu\lambda} Q_S{}^{\nu\lambda} \\ & - \tfrac{1}{2} u_\mu \Phi^{(\text{ext})\mu}{}_{;\nu\lambda} u^\nu u^\lambda Q_S{}^\rho{}_\rho + a_\mu D_\tau \left(\Phi^{(\text{ext})\mu} Q_S{}^\rho{}_\rho \right). \end{aligned} \tag{3.156b}$$

3.B.3 | Scalar self torque

The self torque of a scalar charged body in terms of the renormalized moments is,

$$\begin{aligned} D_\tau S^{\kappa\lambda} \mathcal{P}_\kappa{}^\sigma \mathcal{P}_\lambda{}^\rho = & \mathcal{P}^\sigma{}_\kappa \mathcal{P}^\rho{}_\lambda \Big[2 \Phi^{(\text{ext})[\kappa} Q_S{}^{\lambda]} + 2 \lambda \Phi^{(\text{ext})[\kappa}{}_{;\mu} Q_S{}^{\lambda]\mu} \\ & + \tfrac{2}{3} \lambda q_S D_\tau a^{[\kappa} Q_S{}^{\lambda]} + 2 \lambda \Phi^{(\text{ext})[\kappa} Q_S{}^{\lambda]\mu} a_\mu \Big] + \mathcal{O}(\lambda^2) \end{aligned} \tag{3.157}$$

3.B.4 | Scalar point particle reduced order

We again specialize to a monopole body, for which $S_{\mu\nu} = 0$, $Q_S^\mu = 0$, $Q_S^{\mu\nu} = 0$, and present the reduced order equation of motion.

Here we give the acceleration and rest mass evolution of the point-particle limit for a scalar charge, similar to the expressions for an electromagnetic charge given in Sec. 3.5.4. Note that the lack of a conserved total charge for the scalar case makes this limit somewhat arbitrary - we take it to indicate the vanishing of all moments of the body apart from the renormalized charge q_S.

The acceleration, in terms of only the external field and the charge, is

$$\begin{aligned} a^\kappa =& \frac{q_S}{m}\mathcal{P}^\kappa{}_\sigma\Phi^{(\mathrm{ext})\sigma} + \frac{4}{3}\lambda\frac{{q_S}^2}{m^2}D_\tau q_S\mathcal{P}^\kappa{}_\sigma\Phi^{(\mathrm{ext})\sigma} + \frac{1}{3}\lambda\frac{{q_S}^3}{m^2}\mathcal{P}^\kappa{}_\sigma u^\mu\Phi^{(\mathrm{ext})\sigma}{}_{;\mu} \\ &+ \frac{20}{9}\lambda^2\frac{{q_S}^3}{m^3}\left(D_\tau q_S\right)^2\mathcal{P}^\kappa{}_\sigma\Phi^{(\mathrm{ext})\sigma} + \lambda^2\frac{{q_S}^4}{m^3}\mathcal{P}^\kappa{}_\sigma\left(\frac{10}{9}D_\tau q_S\Phi^{(\mathrm{ext})\sigma}{}_{;\mu}u^\mu + \frac{4}{9}{D_\tau}^2 q_S\Phi^{(\mathrm{ext})\sigma}\right) \\ &+ \frac{1}{9}\lambda^2\frac{{q_S}^5}{m^3}\mathcal{P}^\kappa{}_\lambda\Phi^\lambda{}_{;\mu\nu}u^\mu u^\nu - \frac{4}{9}\lambda^2\frac{q_S^5}{m^4}D_\tau q_S\mathcal{P}^\kappa{}_\sigma\Phi^{(\mathrm{ext})\sigma}u_\lambda\Phi^{(\mathrm{ext})\lambda} \\ &+ \frac{1}{9}\lambda^2\frac{{q_S}^6}{m^4}\big(-\mathcal{P}^\kappa{}_\lambda\Phi^{(\mathrm{ext})\lambda}{}_{;\mu}u^\mu u_\sigma\Phi^{(\mathrm{ext})\sigma} + \mathcal{P}^\kappa{}_\lambda\Phi^{(\mathrm{ext})\lambda}\Phi^{(\mathrm{ext})}{}_{;\mu\nu}u^\mu u^\nu \\ &+ \mathcal{P}^\kappa{}_\sigma\Phi^{(\mathrm{ext})\sigma}{}_{;\mu}\Phi^{(\mathrm{ext})\mu}\big) + \mathcal{O}\left(\lambda^3\right) \end{aligned} \tag{3.158}$$

The evolution of the renormalized mass, in terms of only the external field and the charge, is simply

$$D_\tau m = q_S\Phi^{(\mathrm{ext})\mu}u_\mu + \lambda q_S{D_\tau}^2 q_S + \mathcal{O}(\lambda^3) \tag{3.159}$$

Chapter 3 Bibliography

[152] H. A. Lorentz, *Theory of Electrons* (Dover, 2011, reprinted from 1915)

[153] M. Abraham, Ann. Physik **10** (1903), p. 105

[154] P. A. M. Dirac, *Classical Theory of Radiating Electrons*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **167** (1938) (929), pp. 148–169

[155] T. Erber, *The Classical Theories of Radiation Reaction*, Fortschritte der Physik **9** (1961) (7), pp. 343–392,
URL: http://dx.doi.org/10.1002/prop.19610090702

[156] T. C. Mo and C. H. Papas, *New Equation of Motion for Classical Charged Particles*, Phys. Rev. D **4** (1971), pp. 3566–3571,
URL: https://link.aps.org/doi/10.1103/PhysRevD.4.3566

[157] C. Teitelboim, *Radiation Reaction as a Retarded Self-Interaction*, Phys. Rev. D **4** (1971), pp. 345–347,
URL: https://link.aps.org/doi/10.1103/PhysRevD.4.345

[158] H. Spohn, *Dynamics of charged particles and their radiation field* (1999)

[159] S. E. Gralla, A. I. Harte and R. M. Wald, *A Rigorous Derivation of Electromagnetic Self-force*, Phys. Rev. **D80** (2009), p. 024031

[160] E. Poisson, A. Pound and I. Vega, *The Motion of Point Particles in Curved Spacetime*, Living Reviews in Relativity **14** (2011) (1), p. 7,
URL: https://doi.org/10.12942/lrr-2011-7

[161] LIGO Scientific Collaboration and Virgo Collaboration, *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116** (2016), p. 061102,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.061102

[162] LIGO Scientific Collaboration and Virgo Collaboration, *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, Phys. Rev. Lett. **119** (2017), p. 161101,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.161101

[163] Y. Mino, M. Sasaki and T. Tanaka, *Gravitational radiation reaction to a particle motion*, Phys. Rev. D **55** (1997), pp. 3457–3476,
URL: https://link.aps.org/doi/10.1103/PhysRevD.55.3457

[164] T. C. Quinn and R. M. Wald, *Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime*, Phys. Rev. D **56** (1997), pp. 3381–3394,
URL: http://link.aps.org/doi/10.1103/PhysRevD.56.3381

[165] E. Rosenthal, *Second-order gravitational self-force*, Phys. Rev. D **74** (2006), p. 084018,
URL: http://link.aps.org/doi/10.1103/PhysRevD.74.084018

[166] A. Pound, *Conservative effect of the second-order gravitational self-force on quasicircular orbits in Schwarzschild spacetime*, Phys. Rev. D **90** (2014), p. 084039,
URL: https://link.aps.org/doi/10.1103/PhysRevD.90.084039

[167] A. Pound, *Second-order Gravitational Self-force*, in *Proceedings, 13th Marcel Grossmann Meeting (MG13): Stockholm, Sweden, July 1-7, 2012* (2015), pp. 975–977,
URL: https://inspirehep.net/record/1232382/files/arXiv:1305.1789.pdf

[168] S. Detweiler, *Gravitational radiation reaction and second-order perturbation theory*, Phys. Rev. D **85** (2012), p. 044048,
URL: http://link.aps.org/doi/10.1103/PhysRevD.85.044048

[169] C. R. Galley, *A nonlinear scalar model of extreme mass ratio inspirals in effective field theory: I. Self-force through third order*, Classical and Quantum Gravity **29** (2012) (1), p. 015010, URL: http://stacks.iop.org/0264-9381/29/i=1/a=015010

[170] C. R. Galley, *A nonlinear scalar model of extreme mass ratio inspirals in effective field theory: II. Scalar perturbations and a master source*, Classical and Quantum Gravity **29** (2012) (1), p. 015011,
URL: http://stacks.iop.org/0264-9381/29/i=1/a=015011

[171] B. Wardell and A. Gopakumar, *Self-force: Computational Strategies*, Fund. Theor. Phys. **179** (2015), pp. 487–522

[172] L. Barack, *Gravitational self force in extreme mass-ratio inspirals*, Class. Quant. Grav. **26** (2009), p. 213001

[173] C. R. Galley and B. L. Hu, *Self-force with a stochastic component from radiation reaction of a scalar charge moving in curved spacetime*, Phys. Rev. D **72** (2005), p. 084023,
URL: http://link.aps.org/doi/10.1103/PhysRevD.72.084023

[174] A. I. Harte, *Motion in classical field theories and the foundations of the self-force problem*, Fund. Theor. Phys. **179** (2015), pp. 327–398

[175] A. Pound, *Second-order gravitational self-force*, Phys. Rev. Lett. **109** (2012), p. 051101

[176] S. E. Gralla, *Second Order Gravitational Self Force*, Phys. Rev. **D85** (2012), p. 124011

[177] N. Kumar, K. Z. Hatsagortsyan and C. H. Keitel, *Radiation-Reaction-Force-Induced Nonlinear Mixing of Raman Sidebands of an Ultraintense Laser Pulse in a Plasma*, Phys. Rev. Lett. **111** (2013), p. 105001,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.111.105001

[178] V. I. Berezhiani, S. M. Mahajan and Z. Yoshida, *Plasma acceleration and cooling by strong laser field due to the action of radiation reaction force*, Phys. Rev. E **78** (2008), p. 066403,
URL: https://link.aps.org/doi/10.1103/PhysRevE.78.066403

[179] M. Chen, A. Pukhov, T.-P. Yu and Z.-M. Sheng, *Radiation reaction effects on ion acceleration in laser foil interaction*, Plasma Physics and Controlled Fusion **53** (2011) (1), p. 014004, URL: http://stacks.iop.org/0741-3335/53/i=1/a=014004

[180] J. Krueger and M. Bovyn, *Relativistic Motion of a Charged Particle in a Plane Electromagnetic Wave with Arbitrary Amplitude*, J. Phys. **A9** (1976), pp. 1841–1846

[181] W. G. Dixon, *Dynamics of extended bodies in general relativity. I. Momentum and angular momentum*, Proc. Roy. Soc. Lond. **A314** (1970), pp. 499–527

[182] W. G. Dixon, *Dynamics of extended bodies in general relativity. II. Moments of the charge-current vector*, Proc. Roy. Soc. Lond. **A319** (1970), pp. 509–547

[183] K. Kyrian and O. Semer?k, *Spinning test particles in a Kerr field - II*, Monthly Notices of the Royal Astronomical Society **382** (2007) (4), pp. 1922–1932, URL: http://mnras.oxfordjournals.org/content/382/4/1922.abstract

[184] L. F. O. Costa, J. Natário and M. Zilhão, *Spacetime dynamics of spinning particles: Exact electromagnetic analogies*, Phys. Rev. D **93** (2016), p. 104006, URL: https://link.aps.org/doi/10.1103/PhysRevD.93.104006

[185] A. Pound, *Singular perturbation techniques in the gravitational self-force problem*, Phys. Rev. D **81** (2010), p. 124009

[186] P. D. D'Eath, *Dynamics of a small black hole in a background universe*, Phys. Rev. **D11** (1975), pp. 1387–1403

[187] S. E. Gralla and R. M. Wald, *A Rigorous Derivation of Gravitational Self-force*, Class. Quant. Grav. **25** (2008), p. 205009, [Erratum: Class. Quant. Grav.28,159501(2011)]

[188] *On the classical theory of particles* **194** (1948) (1039), pp. 543–555

[189] *The Classical Theory of Fields* (Oxford: Pergamon press)

[190] L. Parker and J. Z. Simon, *Einstein equation with quantum corrections reduced to second order*, Phys. Rev. **D47** (1993), pp. 1339–1355

[191] E. E. Flanagan and R. M. Wald, *Does back reaction enforce the averaged null energy condition in semiclassical gravity?*, Phys. Rev. **D54** (1996), pp. 6233–6283

[192] F. Rohrlich, *Dynamics of a charged particle*, Phys. Rev. E **77** (2008), p. 046609, URL: https://link.aps.org/doi/10.1103/PhysRevE.77.046609

[193] J. Moxon, http://pages.physics.cornell.edu/~jem497/ (2017), URL: http://pages.physics.cornell.edu/~jem497/

[194] A. Pound, *Gauge and motion in perturbation theory*, Phys. Rev. D **92** (2015), p. 044021, URL: https://link.aps.org/doi/10.1103/PhysRevD.92.044021

[195] E. E. Flanagan and D. A. Nichols, *Conserved charges of the extended Bondi-Metzner-Sachs algebra*, Phys. Rev. D **95** (2017), p. 044002, URL: https://link.aps.org/doi/10.1103/PhysRevD.95.044002

[196] R. M. Wald, *General Relativity* (Chicago Univ. Pr., Chicago, USA, 1984)

[197] T. Quinn, *Axiomatic approach to radiation reaction of scalar point particles in curved space-time*, Phs. Rev. D (2000), p. 62.064029

Chapter 4

The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions, I. Foundations

Coauthors:

Éanna Flanagan, Cornell University

Tanja Hinderer, University of Maryland

Adam Pound, University of Southampton

4.1 | Introduction and summary

The recent direct observations of merging black holes (BHs) by Advanced LIGO [198] have made gravitational waves (GWs) available as a unique new tool for probing the physics and astrophysics of compact objects and strong-field gravity. However, extracting the information contained in such GW signals relies on having accurate theoretical waveform templates for use in matched-filtering data analysis [199]. Computing these templates is a challenging task that requires solving for the dynamical spacetime describing the binary. There are three principal regimes in the parameter space of binary systems, each amenable to different calculational strategies: (i) The weak-field, slow-motion regime describing binaries at large orbital separations. This can be accurately modeled using post-Newtonian (PN) theory, an expansion in the limit $v/c \to 0$ that has been iterated to high order since its first use a century ago [200]. (ii) The relativistic, comparable-mass regime, which must be treated using numerical relativity (NR). Breakthroughs over the last decade have led to numerous successful NR simulations of comparable-mass binary coalescences [201–204], and these

binaries can now be rapidly simulated using a calibrated effective-one-body (EOB) model [205,206] or other composite models [207–210]. (iii) The relativistic, high-mass-ratio regime, where one body is much more massive than the other and the binary evolution is characterized by long, gradual inspirals. This can be modeled using gravitational self-force theory, an expansion in the limit of small mass ratios $\varepsilon = \mu/M$ [211,212]. At zeroth order in this expansion, the smaller body moves as a test particle in the larger body's background spacetime. At subleading orders the smaller body generates a metric perturbation that reacts back on it, modifying the dynamics away from geodesic motion. This high-mass-ratio case is the regime considered in this paper.

The two-body problem in the high-mass-ratio regime has direct observational relevance. So-called intermediate-mass-ratio inspirals (IMRIs) with mass ratios $\sim$ 1:100 could soon be detected by LIGO, Virgo, and KAGRA [213], and they are likely to be better modeled by high-mass-ratio, self-force techniques than by comparable-mass, NR techniques. Moreover, a key source for the planned space-based detector LISA will be extreme-mass-ratio inspirals (EMRIs), in which stellar-mass compact objects ($\mu \sim 1 - 100 M_\odot$) spiral into massive BHs ($M \sim 10^5 - 10^7 M_\odot$) in galactic centers [214–216]. Because its motion is nearly geodesic, the small body lingers in the large BH's strong-field region for numerous ($\sim \varepsilon^{-1}$) GW cycles before merger, thus enabling GW measurements with exquisite accuracy. The GW signals will be rich in information and provide multiple tests of general relativity in this unexplored regime. For example, the signals will encode a detailed map of the spacetime of the massive BH [217,218], allowing precise measurements of the massive BH's parameters and tests of the no-hair theorem. The signals will further contain important astrophysical information about different aspects of stellar dynamics in galactic centers and the growth history of the massive BH [219], and they will potentially yield measurements of the Hubble constant when combined with information about the host galaxy [220,221]. But realizing the science goals for these sources will require meeting a stringent accuracy demand: the waveform models must be accurate to within $\ll 1$ radian over the $\sim \varepsilon^{-1}$ cycles in the highly relativistic regime. This necessitates carrying out the calculations of the gravitational self-force to at least second order in ε [222,223].

At first order, concrete (typically numerical) implementations of self-force theory are thoroughly developed for eccentric, equatorial orbits in both Schwarzschild [224–228] and Kerr back-

grounds [229] and are within reach for generic, inclined orbits. These implementations have been used to compute a litany of physical effects [230–234], have achieved astonishing accuracy [235], and have provided important input to PN theory [236–241], to NR simulations [242], and to EOB methods [243–246]. At second order, implementations are at a much earlier phase of development, but practical formulations exist [247–252] and numerical results have begun to materialize for the simplest case of quasicircular orbits in a Schwarzschild background [253].

However, with few exceptions, these computations have been "snapshots" of binaries on the time scale of a few orbits, during which the deviation from a background geodesic is small. In the familiar, first-order case, these snapshots proceed by fixing the orbit to be a background geodesic, computing the first-order metric perturbation sourced by that geodesic, and from it computing the self-force and related quantities, without allowing the self-force to accelerate the small object. Such a strategy is ideal for the computation of short-term conservative effects, which have been the primary basis for comparisons with PN and NR results. But it leaves the long-term inspiral of the system unaddressed.

Several methods have been devised to account for the binary's long-term evolution. In the original formulation of first-order self-force theory [254, 255], the metric perturbation and worldline of the small object evolve together as a coupled system, in a self-consistent way. This "self-consistent" formalism has since been generalized in a systematic manner to all orders in perturbation theory [256, 257]. In principle, the resulting formulation can be used to directly simulate complete inspirals in the time domain. Unfortunately, such time-domain computations would suffer from cumulative numerical errors and unstable gauge perturbations [258], which would be difficult to cope with over the $\sim 10^5$ orbits in an EMRI. Further, the existing formulation of the self-consistent approximation does not dynamically evolve the mass and spin of the central black hole. Instead, the gauge condition on the initial data surface fixes the first derivative of the central mass for the entire computation. The secular growth of the error in the central mass invalidates the existing self-consistent approximation for radiation-reaction timescales $\sim M^2/\mu$. To date, self-consistent evolutions have only been performed in a scalar model and have been limited to a small number of orbits [259]. More fundamentally, the self-consistent formulation is not ideal because it is not adapted to any particular class of orbits, meaning (a) it does not take advantage of the nearly

periodic nature of a slow inspiral and (b) it does not provide a simple way of estimating the relative contributions of various effects on the specific time scale of an inspiral.

Although direct implementations of the self-consistent formalism remain a goal within the self-force community, most evolution schemes have instead started from the "snapshot" calculations described above. Two such schemes have been explored. One strategy, based on the "method of osculating geodesics" [260, 261], is to compute the self-force at each instant by approximating the source orbit at that instant as the geodesic tangent to the accelerated orbit, effectively evolving from one geodesic to the next. This method has been used to simulate inspirals into a Schwarzschild BH [262,263] using a bank of first-order snapshot computations, and a similar implementation should soon be possible for orbits around a Kerr BH. The second, historically older and more developed evolution scheme is the "adiabatic approximation" [264–268]. It proceeds in a similar manner, using a geodesic source at each instant, but it only incorporates certain time-averaged dissipative effects of the first-order self-force. Because those dissipative effects are generally far easier to compute than the complete first-order force, this method has enabled simulations of generic inspirals in Kerr spacetime, and in restricted cases, of the long-term evolution of the waveform [269–271].

These two schemes have been justified as approximations to the full first-order self-consistent dynamics, and currently they have only been implemented at that order. For accurate modeling, a more general, higher-order approach is now required. Although the osculating-geodesic scheme can be formulated to any order (via an appropriate re-expansion of the self-consistent equations [272]), and similar alternatives have been proposed [273], the practicality of these schemes at nonlinear orders is unclear. However, a more compelling option is available: the method of multiple scales. A multiscale expansion splits the system's dynamics into dependencies on "fast time" and "slow time" variables, corresponding respectively to the orbital time scale $\sim M$ and the inspiral time scale $\sim M/\varepsilon$. By capitalizing on the system's clear separation of time scales and its nearly periodic nature, this method provides a clean, systematic, efficient way of tackling the problem. At each fixed value of slow time, the fast-time equations retain the periodicity of a geodesic, and they can be accurately solved in the frequency domain to obtain snapshots of the system; the slow-time equations then smoothly evolve between those snapshots.

After some preliminary work [274, 275], multiscale methods were first explored as a compre-

hensive approach to high-mass-ratio modelling by two of us in Ref. [222]. There we presented a multiscale expansion of the self-accelerated equation of motion in Kerr spacetime. However, to obtain a complete binary model, that treatment must be combined with an analogous expansion of the Einstein field equations (EFE). Such expansions of the EFE have been considered in the past [276, 277], and specific methods tailored to high-mass-ratio inspirals were recently developed in the context of a scalar toy model in our paper [278]. Our aim here is to take the next step by meshing those methods with the expansion of the equation of motion, thereby completing the program initiated in Ref. [222].

This completed program will provide a systematic way of organizing and attacking high-mass-ratio modeling on both the short and long time scales in the system. But within that broad aim, we also have more specific goals. Our first objective is to develop a practical scheme sufficiently accurate to extract parameters from an observed waveform. The requirements for such an approximation were established in Ref. [222], based on the following uniform expansion of the orbital phase:

$$\phi = \frac{1}{\varepsilon}\left[\phi_0 + \varepsilon\phi_1 + O(\varepsilon^2)\right], \tag{4.1}$$

where the coefficients ϕ_n depend only on the slow-time variable. ϕ_0 can be computed entirely from the time-averaged dissipative piece of the first-order self-force, as in the adiabatic approximation; hence, this is referred to as adiabatic order. ϕ_1 depends on the complete first-order self-force and on the time-averaged dissipative piece of the second-order force; this is referred to as post-adiabatic order. The validity of the expansion (4.1) relies on the assumption that no transient resonance arise during the inspiral, so it applies directly to inspirals in Schwarzschild and circular inclined and elliptic equatorial orbits in Kerr. For generic orbits in Kerr, the orbital frequencies often pass through transient resonances in the respective orbital frequencies [279]. In the case where the orbit is permitted to pass through a resonance during the inspiral requires modification of (4.1) to include a contribution at $\mathcal{O}(\varepsilon^{-1/2})$ that depends directly on the full first-order dissipative self-force and the phase at the onset of the resonance to $\mathcal{O}(\varepsilon^0)$. For the current treatment, we will make the assumption that our approximation is to be used only when no resonance occurs. To keep phase errors small over the course of an inspiral, an evolution scheme must be accurate to post-adiabatic order, and our treatment of the problem will focus on assembling the specific ingredients necessary to achieve

	Adiabatic	Post Adiabatic
Required Order of Self-Force	First-order dissipative	Second-order dissipative + first-order conservative
Errors in Amplitude of Waveform	$\mathcal{O}(\varepsilon)$	$\mathcal{O}(\varepsilon^2)$
Errors in Phase of Waveform	$\mathcal{O}(1)$	$\mathcal{O}(\varepsilon)$

Table 4.1: Characteristics of the adiabatic and post-adiabatic approximations. The post-adiabatic waveform is distinguished by an improved precision and a dependence on higher-order self-force terms.

that goal. For easy reference, we summarize the key features of the adiabatic and post-adiabatic orders in Table 4.1.

Our second specific objective is to provide an unambiguous, intuitive prescription for computing snapshots of an inspiral at second order. Recent synergies between self-force theory and other binary models have generally relied on calculations of conservative effects on the orbital time scale, and doing the same at second order will facilitate further comparisons with NR and PN theory and calibration of EOB and other composite models. As discussed in Ref. [278], cleanly defining snapshots at nonlinear orders is nontrivial because the geodesic approximation, which has been the basis for snapshot computations at first order, leads to pathological infrared singularities at second order. A well-formulated multiscale approximation overcomes this and provides a clear definition of snapshots as pictures of the system at fixed values of slow time.

Regardless of whether one is interested in snapshots or long-term evolution, problems in the infrared are a key feature of nonlinear self-force theory. New features arise not only on long time scales, but also on large spatial scales [278]. These effects confine our multiscale method to a finite region of spacetime, which we refer to as the interaction zone. To surmount the failure of the method outside that zone, we follow Ref. [278] by introducing additional expansions near the system's boundaries. In the far zone, $r \gg M$, we transition to a post-Minkowski expansion, which is then matched to the multiscale solution in a region of mutual validity. This matching at large distances is similar to the strategy commonly employed in PN calculations [200]. Similarly, in the region near the large BH's horizon, we transition to a specialized expansion in the small Boyer-Lindquist radial distance away from the horizon.

Each of these expansions represents a significant undertaking. For that reason, in the present paper we restrict our presentation to two important aspects: laying the foundations of the complete formalism, and showing specifically how it recovers the standard adiabatic approximation at leading order. In a series of followup papers, we will flesh out the details of the method, derive explicit formulas for the second-order field equations in the interaction zone, detail the post-Minkowski and near-horizon expansion, and give explicit prescriptions for complete post-adiabatic computations and second-order snapshots. For pedagogical simplicity, we restrict our attention here to circular or equatorial orbits in Kerr. Generic orbits will involve additional complications due to orbital resonances [223, 280, 281], which introduce a third, transient time scale into the evolution. Inclusion of resonance effects will be the subject of forthcoming work.

The organization of the paper is as follows. We start in Sec. 4.2 with an overview of the computational strategy. Next, we discuss in Sec. 4.3 the analogous systematic multiscale expansion of the EFEs and the remaining gauge freedom. In Sec. 4.4 we briefly review the multiscale treatment of the orbital motion, performing the analysis using coordinate time as the evolution parameter instead of proper time as done previously. The details of the self force require also discussions of the self consistent formalism. Details on the explicit forcing functions that enter the equations of motion are relegated to Appendix 4.A. In Sec. 4.5, we develop a specialization of the geometric optics formalism for portion of the computation distant from the inspiral. In Sec. 4.7 we present the application of our formalism to computing adiabatic-order waveforms. We demonstrate that these waveforms can be computed by making minor modifications to the standard solutions of the Teukolsky equation, that matching to the post-Minkowski and Near-Horizon expansions is unnecessary at adiabatic order, and that the waveforms are gauge-invariant within the relevant class of well-behaved gauges. Finally, Sec. 4.8 contains our conclusions.

4.2 Overview of computational strategy : zones and scales

4.2.1 Foundations and assumptions

We seek to describe the inspiral of a small body of mass μ into a much larger black hole of mass M. Defining the mass ratio $\varepsilon = \mu/M \ll 1$, we consider a one-parameter family of spacetimes[1]

$$(\mathcal{M}, g_{\mu\nu}(\varepsilon), T_{\mu\nu}(\varepsilon)),$$

labeled by ε with $0 < \varepsilon < \varepsilon_0$ for some ε_0, where $T_{\mu\nu}(\varepsilon)$ is the stress-energy tensor of the small body. We make two assumptions:

1. The metric $g_{\mu\nu}$ and stress energy tensor $T_{\mu\nu}$ are jointly smooth in the spacetime coordinates and ε for $\varepsilon > 0$.

2. The Einstein equation is satisfied for all $\varepsilon > 0$

$$G_{\mu\nu}[g(\varepsilon)] = 8\pi T_{\mu\nu}(\varepsilon), \qquad \nabla_\mu(\varepsilon) T^{\mu\nu}(\varepsilon) = 0. \tag{4.2}$$

Consider now one-parameter families $\varphi_\varepsilon : \mathcal{M} \to \mathcal{M}$ of smooth diffeomorphisms which depend smoothly on ε for $\varepsilon > 0$. Such diffeomporphisms can be classified into three types:

(i) those which are independent of ε;

(ii) those which depend on ε and which can be extended continuously to $\varepsilon = 0$;

(iii) those which depend on ε but do not have a continuous limit as $\varepsilon \to 0$.

The assumptions so far are invariant under all three types of transformation, when the metric and stress-energy tensor are transformed by the pullback

$$g_{\mu\nu}(\varepsilon) \to \varphi_{\varepsilon *} g_{\mu\nu}(\varepsilon), \qquad T_{\mu\nu}(\varepsilon) \to \varphi_{\varepsilon *} T_{\mu\nu}(\varepsilon). \tag{4.3}$$

[1] One could also consider more general one parameter families $(\mathcal{M}_\varepsilon, g_{\mu\nu}(\varepsilon), T_{\mu\nu}(\varepsilon))$ where the spacetime manifold depends on ε. However, as long as all the manifolds are diffeomorphic to one another, this does not give rise to any increase in generality, since one can choose an identification of all the manifolds and map the fields via the pullback (4.3) to obtain a physically equivalent one parameter family of solutions.

One parameter families related in this way are physically equivalent. Note however that the one-parameter family of solutions can have very different properties in different gauges when the gauges are related by diffeomorphisms of type (iii). For example, in one gauge the support $\mathcal{T}_\varepsilon$ of the stress energy tensor could have the property that $\mathcal{T}_{\bar\varepsilon} < \mathcal{T}_\varepsilon$ for $\bar\varepsilon < \varepsilon$, and that the intersection of all the supports defines a worldline in spacetime that is independent of ε. In another gauge this might not be true. Or, in one gauge the limit of the metric as $\varepsilon \to 0$ might exist, and in another gauge this might not be true.

We will be considering below a number of different approximation schemes. Each of these will be defined by appending to the list of assumptions above a third assumption of the form

3. On a submanifold $\mathcal{R}_\varepsilon$ the metric $g_{\mu\nu}(\varepsilon)$ and stress energy tensor $T_{\mu\nu}(\varepsilon)$ satisfy the regularity conditions ... and /or have an expansion of the specific form

$$g_{\mu\nu}(\varepsilon) = \ldots, \qquad T_{\mu\nu}(\varepsilon) = \ldots. \tag{4.4}$$

This third assumption is essentially an ansatz which is justified by substituting the form of the assumption into Einstein equations, and showing that solutions to the resulting equations exist and are unique in a suitable sense. In addition this assumption is normally *not* invariant under general ε-dependent diffeomorphisms (although it frequently is invariant under diffeomporphisms of type (i) and type (ii)). The set of such diffeomorphisms which preserve the form of assumption 3 will define a group that will vary from one approximation scheme to another.

For our application to extreme mass ratio inspirals, we will be considering several different approximation schemes of this type, each valid in different regions $\mathcal{R}_\varepsilon$ of spacetime. For each scheme, one solves for the most general possible solution, and the physically appropriate solutions are then obtained by demanding consistency between the solutions of the different schemes within their common domain of validity. This is the method of matched expansions, which has been a standard method for computing the motion of a small body in a background spacetime [212, 254, 272, 276, 282–287]. As mentioned above, the ultimate justification for the assumptions of the form 3 is the fact that one show the existence and uniqueness of solutions obtained starting from these assumptions. In the context of the matched expansion technique, uniqueness is obtained only after

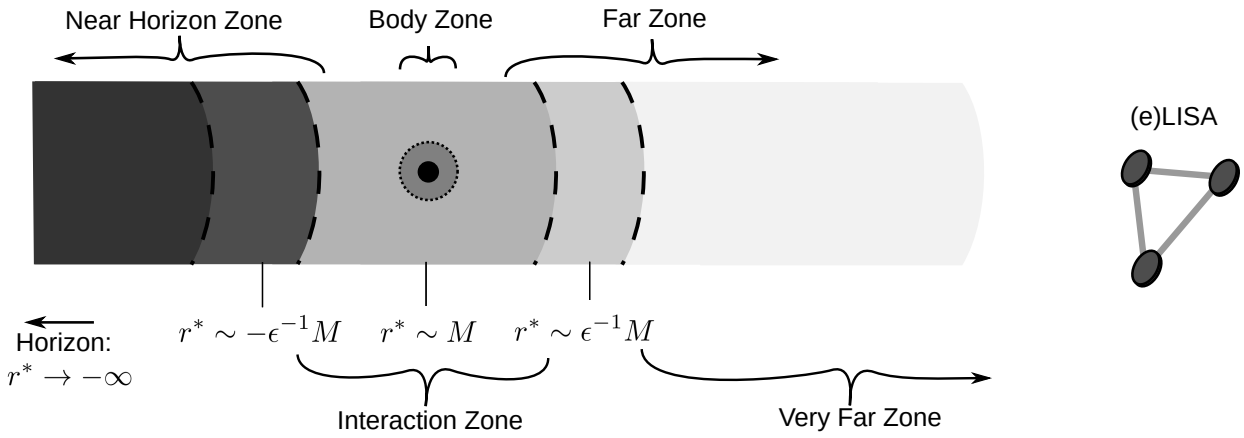

Figure 4.1: An illustration showing the various spacetime regions or zones for a small body of mass $\mu = \varepsilon M$ inspiraling into a black hole of mass M. We use different approximation methods in each zone, which must then be matched consistently together.

all of the matchings have been performed. In this sense, all of the approximation schemes rise or fall together; their justifications are linked.

4.2.2 | Spacetime zones and approximation methods

We next turn to a description of the different regions or zones of spacetime that we consider, and the approximation methods that apply in each zone. The small body of mass μ moves in a background Kerr spacetime of mass M, so combinations of the characteristic scale M of the orbit and the ratios of the masses $\varepsilon = \mu/M$ define the relevant scales of the spacetime. We denote by $\tilde{r}$ the distance from the small body, and by r^* the tortoise radial coordinate of the background spacetime. The various zones we find useful to distinguish in the case of large mass ratio are:

- The *Far Zone* $r^* \gg M$ which is distant from the entire inspiral system. In this region we perform a geometric optics expansion [288, 289]; see Sec. 4.5 below.

- The *Interaction Zone* where $\tilde{r} \gg \varepsilon M$ and $-\varepsilon^{-1}M \ll r^* \ll \varepsilon^{-1}M$. In this region we will use the self-consistent expansion and the multiscale expansion [222, 274–276, 278]; see Secs. 4.3 below.

- The *Body Zone* or *Puncture Zone* where $\tilde{r} \ll M$. In this region we will use the self-consistent expansion [212, 256, 272]; see Sec. 4.4 below.

- The *Near Horizon Zone* $-r^* \gg M$ very close to the black hole horizon. In this region we use a near horizon expansion, which will be presented in detail in the forthcoming publication [290].

There are three different matchings necessary for determine the global solution: a matching between the body zone and the interaction zone (which on short timescales is the usual matching used to derive the self force), a matching between the interaction zone and the far zone [222, 278], and a matching between the interaction zone and the near horizon zone [278]. In both the Near-Horizon Zone and the Far Zone, the first order field equations will give only homogeneous modes, which will be fixed by matching to the first order solution in the interaction zone. Post-adiabatic calculations will require accuracy in the metric perturbation for these outer regions through $\mathcal{O}(\varepsilon^2)$, and the matching will impose nontrivial constraints on the interaction zone. The analogous scalar derivation was presented by [278], in which a similar effect was computed. Our derivation of the post-adiabatic effects will be presented in our future paper [291]. In addition, it is worthwhile to introduce the sub-zone of the Far Zone: the *Secular Far Zone* $r^* \gg M/\varepsilon$, at which the retardation effects associated with secular evolution of the radiation can no longer be neglected; see also 4.5 below.

For the matching to proceed smoothly, we find it necessary to introduce a new time variable,

$$w = t + h(r^*), \tag{4.5}$$

such that the slow evolution of the system will no longer be specified as slices of constant Boyer-Lindquist time t. We impose that

$$\lim_{r^* \to \infty} h(r^*) = -r^*, \tag{4.6a}$$

$$\lim_{r^* \to -\infty} h(r^*) = r^*, \tag{4.6b}$$

such that w asymptotically approaches the advanced time v at the horizon and the retarded time u at future null infinity. Further, we also impose that these convergences occur sufficiently quickly that

$$r^* \sim M/\varepsilon \Rightarrow w - u \sim \varepsilon M, \tag{4.7a}$$

$$-r^* \sim M/\varepsilon \Rightarrow w - v \sim \varepsilon M \tag{4.7b}$$

In addition, we define a normal to hypersurfaces Σ_w of constant w, which we denote w^μ. To keep the computation general to various convenient choices for $h(r^*)$, we do not impose any particular function which satisfies the above conditions, and instead carry through the h dependence in our computations. In appendix 4.B we give a more expanded discussion of the use of the adjusted time variable.

In the following collection of sections, we explain the full details of the approximation methods in each of the distinct spacetime regions.

4.3 The multiscale approximation in the Interaction Zone

4.3.1 Preamble

Black hole perturbation theory is a natural approximation method for use in the Interaction Zone, which is chosen to encompass the strong-field region of the large companion $|r^*| \ll M/\varepsilon$, but to exclude the strong-field region of the small companion $\tilde{r} \gg \varepsilon M$. In this region, the additional metric perturbations sourced by the small companion are perturbatively small $\sim \varepsilon$ compared to the background metric of the large companion. Locally in time, black hole perturbation theory is robust.

It is well-known, however, that the precise computation of the state of the system develops $\mathcal{O}(1)$ errors after a dephasing time $\sim M/\sqrt{\varepsilon}$, which is inadequate for a computation which hopes to describe the full inspiral lasting for radiation-reaction timescale $\sim M/\varepsilon$ [222,273]. The principal benefits of using a multiscale method have been previously discussed in Refs. [222,274,275,278], and here we review the case for multiscale techniques. The key idea in the method is an ansatz for the ε dependence for one-parameter family of metric perturbations $g_{\alpha\beta}(\varepsilon)$ and worldlines $z^\mu(\varepsilon)$ that is well-tuned to the dynamical system at hand and is justified *a posteriori* via order by order solution of the Einstein field equation. The full development of the multiscale approximation for orbital motion of $z^\mu(\varepsilon)$ through post-adiabatic order was presented in [222]. In this section, we describe the extension of the multiscale approximation to the field equations in the Interaction Zone.

The critical piece of information provided by the multiscale analysis of the orbital motion is the gradual evolution of the frequencies Ω^r, Ω^θ, and Ω^ϕ associated with the r,θ, and ϕ components of

the near-geodesic motion. We write the slowly evolving frequencies in terms of a scaled 'slow' time,

$$\tilde{w} \equiv \varepsilon w, \tag{4.8}$$

so that the frequencies take form,

$$\Omega^A = \Omega^A(\tilde{w}, \varepsilon) = \Omega^{A(0)}(\tilde{w}) + \varepsilon \Omega^{A(1)}(\tilde{w}) + \mathcal{O}(\varepsilon^2). \tag{4.9}$$

For this set of frequencies we define a collection of phase variables $\varphi^A = (\varphi^r, \varphi^\theta, \varphi^\phi)$,

$$\varphi^A(\tilde{w}, \varepsilon) \equiv \varphi_0^A + \frac{1}{\varepsilon} \int_0^{\tilde{w}} d\tilde{w}' \Omega^A(\tilde{w}', \varepsilon), \tag{4.10}$$

for some set of initial phases φ_0^A

For any such definition of time variables (4.10), (4.8), it holds that if any function $f'(x^i, \varphi^A, \tilde{w})$ is a solution to a differential equation on $(\varphi^A, \tilde{t}, x^i)$,

$$D'[f'] = S(x^i, \varphi, \tilde{w}), \tag{4.11}$$

where

$$D' = D|_{\partial t \to \Omega^A \partial_{\varphi^A} + \varepsilon \partial_{\tilde{w}}}, \tag{4.12}$$

and D is any differential operator, then $f(x^i, w) = f'(x^i, \varphi^A(\varepsilon w), \varepsilon w)$ satisfies

$$D[f] = S(x^i, \varphi^A(\varepsilon w), \varepsilon w) \tag{4.13}$$

For the remainder of the paper, we will omit the prime on such differential operators, taking the replacement (4.12) wherever a differential operator acts on a function of the multiscale variables $x^i, \tilde{w}, \varphi^A$.

4.3.2 | Ansatz for the metric

We assume a one-parameter family of solutions of Einsteins equation satisfying assumptions 1 and 2 of Sec. 4.2.1 above. In the Interaction Zone, we assume that for both the worldline $z^\alpha(x^\mu)$ and the metric $g_{\alpha\beta}(x^\mu)$, there exists additional functions $z'^\alpha(x^i, \tilde{t}, \varphi^A)$ and $g'_{\alpha\beta}(x^i, \tilde{t}, \varphi^A)$, such that z'^α and $g'_{\alpha\beta}$ are multiperiodic in φ^A. We assume also that there exist functions $\varphi^A(t), \tilde{t}(t)$ satisfying

(4.8) and (4.10), such that

$$z'^{\alpha}(x^i, \varphi^A(t), \tilde{t}(t), \varepsilon) = z^{\alpha}(x^i, t, \varepsilon) \tag{4.14a}$$

$$g'_{\alpha\beta}(x^i, \varphi^A(t), \tilde{t}(t), \varepsilon) = g_{\alpha\beta}(x^i, t, \varepsilon). \tag{4.14b}$$

For notational convenience, we will omit the primes for these promoted functions and simply use the function arguments to distinguish between $\{z, g\}$ and $\{z', g'\}$.

Additionally, we assume that there exist some constants φ_0^A and a set of gauges in the Interaction Zone in which the metric is of the form

$$g_{\alpha\beta}(w, x^i, \varepsilon) = g^{(0)}_{\alpha\beta}(x^i) + \varepsilon h^{(1)}_{\alpha\beta}(\varphi^A, \tilde{w}, x^i) + \varepsilon^2 h^{(2)}_{\alpha\beta}(\varphi^A, \tilde{w}, x^i) + O(\varepsilon^3), \tag{4.15}$$

where $g^{(0)}_{\alpha\beta}$ is the background Kerr metric. On the right hand side of Eq. (4.15), the $O(\varepsilon^3)$ refers to an asymptotic expansion associated with the limit $\varepsilon \to 0$ at fixed φ^A, x^i, and $\tilde{w}$ that is uniform in $\tilde{w}$. Finally, the functions $h^{(1)}_{\alpha\beta}$ and $h^{(2)}_{\alpha\beta}$ are assumed to be multiply periodic in φ^r, φ^θ and φ^ϕ with period 2π in each variable.

4.3.3 | Discussion

As stated in the introduction, we restrict our discussion in this paper to inspirals that do not pass through orbital resonances. The frequencies are determined by a matching computation to the Body Zone which determines the small companion's worldline. The physical interpretation is that φ^A are the values of the orbit's angle variables at the intersection of the inspiraling orbit with the hypersurface $w =$ constant.

We develop the multiscale expansion as a method both for permitting a simplification closely analogous to osculating geodesics at leading order, and for computing the delicate second-order effects important for EMRI detection and parameter estimates and realizing the LISA mission's science goals [292].

The osculating geodesics approximation [260] simplifies EMRI computations by finding instantaneous self force values by approximating the orbit instantaneously as a geodesic. The relative error in this assumption scales with the mass ratio, so it can yield acceptable results for an adiabatic order computation, which is sufficient to find the amplitude and frequencies of an asymptotic waveform up to $\mathcal{O}(\varepsilon)$ error.

The long inspiral causes the small $\mathcal{O}(\varepsilon)$ error in the frequencies to build secularly to an $\mathcal{O}(1)$ error in the phase of the waveform, which would be bothersome for matched filtering procedures. We must compute the orbital frequencies with only $\mathcal{O}(\varepsilon^2)$ error to compute the waveform phase with $\mathcal{O}(\varepsilon)$ error.

The main advantage of the multiscale technique is that it permits a Fourier decomposition of the relevant physical quantities, such that the coefficients of the decomposition vary slowly with time. This type of behavior is precisely what we'd like to capture for a slow inspiral characteristic of extreme mass ratios. More complete details of the multiscale construction, and example applications are given in [293].

To post-adiabatic order, we assume that we may write the metric and the worldline as the ansatz:

$$g_{\mu\nu}(x^\mu) = g^{(0)}_{\mu\nu}(x^i) + \varepsilon h^{(1)}_{\mu\nu}(\varphi^A, \tilde{w}, x^i) + \frac{1}{2} h^{(2)}_{\mu\nu}(\varphi^A, \tilde{w}, x^i) + \mathcal{O}(\varepsilon^3) \tag{4.16}$$

$$z^\mu = z^{(0)\mu}(\varphi^A, \tilde{w}) + \varepsilon z^{(1)\mu}(\varphi^A, \tilde{w}) + \mathcal{O}(\varepsilon^2) \tag{4.17}$$

This expansion is somewhat different from the metric expansion used in the self-consistent formalism. When (4.16) is evaluated at the physical $\{\tilde{w}(w), \varphi^A(w)\}$, the leading order metric perturbation can be written as a functional only of the leading order worldline $z^{(0)\mu}(\tilde{w}(w), \varphi^A(w))$. Additionally, any time-dependent, non-oscillatory ("secular") contributions are parameterized by the scaled time $\tilde{w}$. The effect of the reparameterization and the adjustment of the differential equation (4.12) is that secular results are promoted by a factor of $1/\varepsilon$, due to build-up over the long inspiral.

A concrete example of this order promotion is the mass of the central black hole. Instantaneously, the deviation from the original black hole mass is a second order $\mathcal{O}(\varepsilon^2)$ effect. However, the total accreted mass from gravitational radiation scales with the duration of the inspiral, so is promoted by a factor of $1/\varepsilon$ to a first-order effect.

The puncture metric $h^{\mathcal{P}}_{\mu\nu}$ derived in [250] was found using the self consistent formulation, which expands the metric at fixed worldline γ (4.51). Due to our use of a perturbatively expanded worldline (4.17), we require the expression for an adjusted $h^{(2)\mathcal{P}}_{\mu\nu}$. At post-adiabatic order, the puncture metric will obtain a time-dependent dipole correction associated with $z^{(1)\mu}(\tilde{w}, q^A)$.

4.3.4 Limited domain of validity of the multiscale approximation

The multiscale formalism derives the metric perturbation by solving order-by-order the wave equation associated only with the phase variables φ^A

$$h^{(n)}_{\mu\nu} = \left(\Box^{-1}_{\varphi} S^{(n)}\right)_{\mu\nu}, \tag{4.18}$$

and by determining the slow time $\tilde{w}$ dependence by imposing fast time φ^A periodicity.

However, if we attempt to use the multiscale formalism to describe the entire relevant region to the the EMRI, we must solve the field equations for the region within $r \sim M/\varepsilon$, due to the outward propagating waves which act as a source for second order metric perturbations.

As was found in [278], the scalar analog to the multiscale computation fails to have a finite second-order solution when the domain of integration for (4.18) is taken over the full space. The second order source must be integrated against the Green function for the spacetime, over the full past lightcone. In a physical solution, this quantity will decay in the distant past, as the inspiral is not eternal and does not distribute infinite energy. However, the multiscale formalism constructs the fixed-$\tilde{w}$ solution to be perfectly periodic, so integrals which act only on the periodic dependence will diverge if taken over a global domain.

The dominant part of the quadratic source arises from the beating of two oscillatory components with the same frequency, each of which has a leading order dependence at large scales of $1/r$. The dominant part of the source then scales as $S^{(2)} \sim \Omega^2/r^2$, and has no oscillatory dependence. This source, integrated against the Green's function, gives rise to a divergent second order field, dramatically signaling the breakdown of the Multiscale formalism when applied globally.

As was also suggested in [278], The solution to this problem is to limit the computational domain of the multiscale formalism to the Interaction zone $-M/\varepsilon \ll r^* \ll M/\varepsilon$, and perform distinct computations in the surrounding regions. The computations in the Near Horizon $-r^* \gg M$ and Far Zones $r^* \gg M$ are then constrained to obey matching conditions in common regions of validity.

4.3.5 Multiscale expansion of the Einstein field equations

To obtain the general form of the metric perturbations, we note that the linearized Einstein operator $G_{\mu\nu}$ does not have any explicit time dependence, so the time dependence of the solutions must derive from the time dependence of the source. The particular source we consider is a particle of mass μ on a worldline $z^\mu(\varphi^A, \tilde{w})$. It is shown in [222], and reviewed in section 4.4, that to post-adiabatic order, the frequency variables φ^A may be directly identified with the action-angle variables associated with the near-geodesic orbital motion q^A. From the multiscale treatment of the orbit discussed in [222] it follows that the source terms in the Puncture Zone have the form

$$T^{(s)}_{\mu\nu}(x^j, w) = T^{(s)}_{\mu\nu}(x^j, q^A(w), \tilde{w}(w)). \tag{4.19}$$

Within the interaction zone, our ansatz for the form of the metric with the perturbations sourced by (4.19) is then Eq. (4.16).

We next use the expansions of the source (4.19) and the metric perturbation (4.16) in the Einstein field equations to derive the perturbative field equations satisfied by the order-by-order expansion of the metric. We expand the operators appearing in the expansion of the Einstein field equations and read off the Einstein equation at each order in ε at fixed $(x^i, q^A, \tilde{w})$. We first define the perturbative expansion of the Einstein field operator in powers of the metric perturbation as,

$$G_{\mu\nu}[g^{(0)} + \varepsilon h] = G_{\mu\nu} + \varepsilon\delta G_{\mu\nu}[h] + \varepsilon^2\delta^2 G_{\mu\nu}[h,h] + \varepsilon^3\delta^3 G_{\mu\nu}[h,h,h] + \mathcal{O}(\varepsilon^4). \tag{4.20}$$

In addition, we denote with a superscript the explicit number of powers of ε which arise from the expansion of the derivatives as (4.12). Then, for each order in the differential operator expansion (4.20), we have the re-expansion,

$$\delta^n G_{\mu\nu}[h\ldots] = \delta^n G^{(0)}_{\mu\nu}[h\ldots] + \varepsilon\delta^n G^{(1)}_{\mu\nu}[h\ldots] + \varepsilon^2 G^{(2)}_{\mu\nu}[h\ldots] + \mathcal{O}(\varepsilon^3), \tag{4.21}$$

where the additional powers of ε arise from the derivative replacements

$$\partial_w \to \Omega^{A(0)}\partial_{q^A} + \varepsilon\left(\partial_{\tilde{w}} + \Omega^{A(1)}\partial_{q^A}\right) + \varepsilon^2\Omega^{A(2)}\partial_{q^A} + \mathcal{O}(\varepsilon^3). \tag{4.22}$$

The full expansion of the Einstein field equations in the Interaction Zone (for which $T_{\mu\nu}$ explicitly vanishes) therefore reads,

$$0 = \varepsilon\delta G^{(0)}_{\mu\nu}[h^{(1)}] + \varepsilon^2\left(\delta G^{(0)}_{\mu\nu}[h^{(2)}] + \delta^2 G^{(0)}_{\mu\nu}[h,h] + \delta G^{(1)}_{\mu\nu}[h^{(1)}]\right) + \mathcal{O}(\varepsilon^3). \tag{4.23}$$

Then, the first and second order equations to be solved purely in the Interaction Zone are,

$$\delta G^{(0)}_{\mu\nu}[h^{(1)}] = 0 \tag{4.24}$$

$$\delta G^{(0)}_{\mu\nu}[h^{(2)}] = -\delta^2 G^{(0)}_{\mu\nu}[h,h] - \delta G^{(1)}_{\mu\nu}[h^{(1)}]. \tag{4.25}$$

In other words, the leading metric perturbation $h^{(1)}$ is precisely a homogeneous solution in the fast time variables, and the subleading metric perturbation $h^{(2)}$ acquires source terms both from the ordinary quadratic source from black hole perturbation theory and from the subleading contributions due to the ε dependence of the multiscale time variables. The slow-time dependence of the metric perturbations are set by a combination of consistency conditions with other zones and with the multiscale approximation method itself:

1. The slow time dependence associated with secular evolution of the inspiral, parameterized by $(E(\tilde{w}), L_z(\tilde{w}), Q(\tilde{w}))$ is enforced by matching with the Body Zone. See section 4.4, 5.2, and discussion below.

2. The slow time dependence of the mass and spin of the central body will enter as a post-adiabatic effect, and are fixed by the fast time and physical angle average of the second-order multiscale Einstein field equation. See Section 5.3.5.

3. Further slow time dependence of the spacetime is pure gauge at post-adiabatic order, and under a suitable choice of gauge, is strictly of post-2-adiabatic order, so is neglected in this series of papers.

In practical computations, the Einstein field equations will likely not be solved separately in the Interaction Zone and the Body Zone. Modern treatments instead opt for an 'effective source' method, or other alternative method for implicitly regularizing the divergent source term at the point of the small companion, while treating the entire strong-field region in the same computation. In Chapter 5, we present the full details under the construction of the effective source method, which is the most convenient construction for our current discussion, although other methods may be used.

In an adiabatic-order evolution, we neglect all effects of post-adiabatic order including conservative first order self force and dissipative second order self force. For adiabatic order, it is sufficient

to solve either the Lorenz-gauge wave equation for a pointlike source,

$$\nabla_\mu \bar{h}^{(1)\mu\nu} = 0 \tag{4.26a}$$

$$E^{(1,0)}_{\mu\nu}[h^{(1)}] = \bar{T}^{(1)}_{\mu\nu}, \tag{4.26b}$$

or the corresponding Teukolsky equation for a pointlike source. For brevity, a full description of the Teukolsky formalism is omitted in this chapter, and reserved for a full discussion of post-adiabatic order computations presented in Chapter 5.

Note that this is a purely periodic form of the gauge-fixed wave equation. For any $\tilde{w}$, this equation has a solution that corresponds to the metric sourced by a geodesic as a function of q^A. The correct $\tilde{w}$ dependence is fixed by the orbital equations (4.76). Using the expansion (4.19) in the condition for stress-energy conservation $\nabla_\nu T^{\mu\nu} = 0$, it follows that the right hand side of (4.26b) is conserved with respect to the background derivative operator with only fast-time derivatives $\nabla^{(0)}_\mu$.

4.3.6 | Gauge transformations

Consider a one-parameter family of spacetimes $(\mathcal{M}, g_{ab}(\varepsilon))$. Then $(\mathcal{M}, \chi^*_\varepsilon g_{ab}(\varepsilon))$ represents the same physical one-parameter family, where χ_ε is an arbitrary one-parameter family of diffeomorphisms $\chi_\varepsilon : \mathcal{M} \to \mathcal{M}$ with the identity map corresponding to $\varepsilon = 0$. We can express such a map as

$$\chi_\varepsilon = \chi_0 \circ \mathcal{D}_{\xi^{(1)}}(\varepsilon) \circ \mathcal{D}_{\xi^{(2)}}(\varepsilon^2/2) \circ \ldots \circ \mathcal{D}_{\xi^{(n)}}(\varepsilon^n/n!), \tag{4.27}$$

where for any vector field ξ^a the flow $\mathcal{D}_\xi(\varepsilon) : \mathcal{M} \to \mathcal{M}$ denotes the one-parameter group of diffeomorphisms generated by ξ^a. Such a map χ is called a knight diffeomorphism, by inspiration from chess moves, because it displaces a point of $\mathcal{M}$ a parameter interval ε along the integral curve of $\xi_{(1)}$, then an interval $\varepsilon^2/2$ along the integral curve of $\xi_{(2)}$, etc. Bruni et al. [294] have shown that any one-parameter family of diffeomorphisms can always be expanded in this way.

The expansion around $\varepsilon = 0$ of the pull-back associated with χ is given by:

$$\begin{aligned} \chi^*_\varepsilon T &= \varphi^*_{\xi_{(1)}}(\varepsilon) \ldots \varphi^*_{\xi_{(k)}}(\varepsilon^k/k!) T \\ &= \sum_{l_1=0}^{\infty} \cdots \sum_{l_k=0}^{\infty} \frac{\varepsilon^{l_1+2l_2+\ldots+kl_k}}{2^{l_2} \ldots (k!)^{l_k} l_1! \ldots l_k!} \chi_0 \circ \pounds^{l_1}_{\xi_{(1)}} \ldots \pounds^{l_k}_{\xi_{(k)}} T \end{aligned} \tag{4.28}$$

where T is any tensor field and $\pounds_\xi$ is the Lie derivative along ξ^a. We assume that a perturbed tensor field $T(\mathbf{x}, t)$ can be expanded around $\varepsilon = 0$ as

$$T(\varepsilon) = T^{(0)} + \sum_{s=1}^{\infty} T^{(s)} \varepsilon^s \tag{4.29}$$

Truncating the expansion (4.27) of the map χ at $O(\varepsilon^2)$, and using (4.28) gives the transformation properties of the expansion coefficients of the tensor fields T and $\bar{T}(\varepsilon) = \chi_\varepsilon^* T$ in two different gauges in the special case of $\chi_0 = \mathbb{1}$ as

$$\bar{T}(\varepsilon) = T^{(0)} + \varepsilon \left(T^{(1)} + \pounds_{\xi_{(1)}} T^{(0)} \right) + \varepsilon^2 \left[T^{(2)} + \left(\pounds_{\xi_{(2)}} + \pounds^2_{\xi_{(1)}} \right) T^{(0)} + 2 \pounds_{\xi_{(1)}} T^{(1)}, \right] \tag{4.30}$$

In what follows, we briefly discuss how these results need to be modified to mesh with the two-time expansion, restricting the discussion to transformations that leave the time-like Killing vector unaffected. Adapting the transformations requires gauge vectors that are consistent with the use of two times and specializing the Lie derivative to the expansion (4.12) of the derivative operator. As before, the multiscale formalism changes the mathematical meaning of the expansions to be $\varepsilon \to 0$ at fixed q^A, $\tilde{w}$.

The multiscale expansion introduces an expanded set of degrees of freedom $\{q^A, \tilde{w}\}$, but with a periodicity restriction. As a result, the set of gauge transformations may alter any of the coordinates $\{x^i, q^A, \tilde{w}\}$, but only in such a way that the form of the multiscale ansatz (4.16) preserved. We also impose at first order that the gauge be chosen such that i) at fixed $\tilde{w}$, $\lim_{r\to\infty} h_{\mu\nu} \to 0$, and ii) $h_{\mu\nu}$ is continuous everywhere. In what follows, we consider the specialization to $\chi_0 = \mathbb{1}$, and we discuss the possibility of $\mathcal{O}(1)$ diffeomorphisms which nonetheless preserve the background metric in the multiscale formalism from Section 4.3.2.

We denote the set of valid transformations

$$\xi^\mu = \varepsilon \xi^{(1)\mu}(x^i, q^A, \tilde{w}) + \varepsilon^2 \xi^{(2)\mu}(x^i, q^A, \tilde{w}) + \mathcal{O}(\varepsilon^3) \tag{4.31a}$$

$$\zeta^A = \varepsilon \zeta^{(1)A}(q^A, \tilde{w}) + \varepsilon^2 \zeta^{(2)A}(q^A, \tilde{w}) + \mathcal{O}(\varepsilon^3), \tag{4.31b}$$

where the expansion of a tensor in distinct gauges is given as

$$\begin{aligned}\bar{T}(\varepsilon) = T^{(0)} &+ \varepsilon\left(T^{(1)} + \pounds^{(0)}_{\xi^{(1)}} T^{(0)} + \zeta^{(1)A}\partial_{q^A} T^{(0)}\right) \\ &+ \varepsilon^2\Big(T^{(2)} + \pounds^{(1)}_{\xi^{(1)}} T^{(0)} + \pounds^{(0)}_{\xi^{(2)}} T^{(0)} + \zeta^{(2)A}\partial_{q^A} T^{(0)} \\ &\qquad + \frac{1}{2}\left(\zeta^{(1)A}\partial_A + \pounds^{(0)}_{\xi^{(1)}}\right)^2 T^{(0)} + \left(\zeta^{(1)A}\partial_{q^A} + \pounds^{(0)}_{\xi^{(1)}}\right) T^{(1)}\Big).\end{aligned} \tag{4.32}$$

Here, the superscript on the Lie derivative describes the contributions which act on $\tilde{w}$ and those that act on q^A with a coefficient $\Omega^{(1)}$, as in the superscript of the Einstein field operator (4.21).

For the metric (4.16), this gauge transformation simplifies due to the purely spatial dependence of the background metric.

$$\begin{aligned}\bar{g}(\varepsilon) = g^{(0)} &+ \varepsilon\left(g^{(1)} + \pounds^{(0)}_{\xi^{(1)}} g^{(0)}\right) \\ &+ \varepsilon^2\left(g^{(2)} + \pounds^{(0)}_{\xi^{(2)}} g^{(0)} + \frac{1}{2}\left(\pounds^{(0)}_{\xi^{(1)}}\right)^2 g^{(0)} + \left(\zeta^{(1)A}\partial_{q^A} + \pounds^{(0)}_{\xi^{(1)}}\right) g^{(1)}\right).\end{aligned} \tag{4.33}$$

The extra degrees of gauge freedom in ζ^A describe a small, slowly varying phase shift. This freedom may be used to perform a near-identity transformation [293, 295, 296], which can greatly simplify the orbital equations of motion. We discuss the use of the near-identity transformation in the multiscale context in appendix 5.A.

Consider also the effect of the gauge restriction on the transformations of the self force. A general gauge vector, as a function of (x,w) can be used to entirely remove the effect of the first-order self force [297, 298]. The permissible class of gauges is constrained in the multiscale approximation. We insist that our metric ansatz (4.16) be preserved by the gauge transformation, which restricts the gauge vector to be of the form $\xi(x, q^i, \tilde{w})$, and to depend periodically on q^i. This restriction ensures that there are portions of the first-order self force that cannot be gauged away.

In particular, the q^i-averaged derivatives of the orbital quantities are preserved under a gauge transformation. The energy and the angular momentum are expressed as:

$$\left\langle \frac{dE}{dw} \right\rangle = \left\langle \frac{d\tau}{dw} \xi_\mu a^\mu \right\rangle \tag{4.34}$$

Using the explicit form of the first order self force, and discarding time averages of purely oscillatory values, this can be simplified to

$$\left\langle \frac{d\tau}{dw} \xi_\mu a^\mu \right\rangle = \left\langle \frac{1}{2}\frac{d\tau}{dw} u^\mu u^\nu \mathcal{L}_\xi h_{\mu\nu} \right\rangle, \tag{4.35}$$

for which the gauge transformation produces a correction $\langle(1/2)d_t(\xi_{\alpha,w}u^\alpha)\rangle$. This demonstrates the component of the first-order self force that is gauge-invariant for the multiscale

The same effect can be seen by applying the gauge transformation to the worldline (4.17), the effect of the restriction is apparent. The gauge-transformed worldline takes the form:

$$z'^\mu = z^{(0)\mu}(P^i(\tilde{w}), q^i) + \varepsilon(z^{(1)\mu}(\tilde{w}, q^i) + \xi^\mu(\tilde{w}, q^i)) + \mathcal{O}(\varepsilon^2) \tag{4.36}$$

The ξ may be used to gauge away some of the adjusted motion, but is restricted from having the necessary large adjustment to remove the $\tilde{w}$ dependence from $z^{(0)}$.

4.3.6.1 Ambiguity in the decomposition of the multiscale gauge and relaxed Einstein field equation

There is an ambiguity at post-adiabatic order in the generalization of any gauge condition which depends on time derivatives to an analogous multiscale gauge. In particular, there are at least two separate multiscale gauges that can both be said to be natural extensions of the Lorenz gauge. The first natural form of the Lorenz gauge equations of motion is

$$0 = \nabla^{(0)}_\mu \bar{h}^{(2)\mu\nu} + \nabla^{(1)}_\mu \bar{h}^{(1)\mu\nu}, \tag{4.37a}$$

$$E^{(0)}_{\mu\nu}[\bar{h}^{(2)}] = \delta^2 R^{(0)}_{\mu\nu}[\bar{h}^{(1)}, \bar{h}^{(1)}] - E^{(1)}_{\mu\nu}[\bar{h}^{(1)}]. \tag{4.37b}$$

An alternative formulation has no multiscale derivatives in the gauge condition, and takes the form,

$$0 = \nabla^{(0)}_\mu \bar{h}'^{(2)\mu\nu} \tag{4.38a}$$

$$E^{(0)}_{\mu\nu}[\bar{h}'^{(2)}] = \delta^2 R^{(0)}_{\mu\nu}[\bar{h}'^{(1)}, \bar{h}'^{(1)}] - R^{(1)}_{\mu\nu}[\bar{h}'^{(1)}] \tag{4.38b}$$

We now demonstrate explicitly that there exists a gauge transformation that satisfies the multiscale construction, which transforms between the first form of the equations of motion (4.37) and the second (4.38). Consider a gauge transformation ξ which adjusts the gauge condition above, but preserves the first order gauge condition

$$\nabla^{(0)}_\mu \bar{h}^{(1)\mu\nu} = \nabla^{(0)}_\mu \bar{h}'^{(1)\mu\nu} = 0. \tag{4.39}$$

We therefore consider a second order gauge transformations $\xi^{(2)}$, for which

$$\bar{h}^{(2)}_{\mu\nu} - \bar{h}'^{(2)}_{\mu\nu} = \nabla^{(0)}_{\mu}\xi^{(2)}_{\nu} + \nabla^{(0)}_{\nu}\xi^{(2)}_{\mu} - g^{(0)}_{\mu\nu}\nabla^{(0)\lambda}\xi^{(2)}_{\lambda} \tag{4.40}$$

Then, the difference in the gauge conditions for the two forms of the wave equations (4.37) and (4.38) offer the constraint,

$$\begin{aligned}\nabla^{(0)}_{\mu}(\bar{h}^{(2)\mu\nu} - \bar{h}'^{(2)\mu\nu}) =& -\nabla^{(1)}_{\mu}\bar{h}^{(1)\mu\nu} \\ \Box^{(0)}\xi^{\nu}(\tilde{w}, q^A, x^i) =& -w_{\mu}(\partial_{\tilde{w}}h^{(1)\mu\nu}(\tilde{w}, q^A, x^i) + \Omega^{(1)A}\partial_A h^{(1)\mu\nu}),\end{aligned} \tag{4.41}$$

which is an inhomogeneous wave equation, solvable with the standard Green's function treatment.

Finally, we verify that this alteration also makes the anticipated change to the relaxed Einstein Field equation. The difference on the right hand side of the relaxed EFE's (4.37),(4.38),

$$\begin{aligned}E^{(1)}_{\mu\nu}[\bar{h}^{(1)}] - R^{(1)}_{\mu\nu}[\bar{h}^{(1)}] =& -g^{(0)}_{\mu\nu}\nabla^{(0)}_{\lambda}w_{\sigma}(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_A)\bar{h}^{(1)\lambda\sigma} + \nabla^{(0)}_{\nu}\left(w_{\lambda}(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_A)\bar{h}^{(1)\lambda}_{\mu}\right) \\ &+ \nabla_{\mu}\left(w_{\lambda}(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_A)\bar{h}^{(1)\lambda}_{\nu}\right),\end{aligned} \tag{4.42}$$

where the difference in the left hand side due to the suggested gauge transformation is

$$\begin{aligned}E^{(0)}_{\mu\nu}[\bar{h}^{(2)}] - E^{(0)}_{\mu\nu}[\bar{h}'^{(2)}] =& \Box^{(0)}(\nabla^{(0)}_{\mu}\xi^{(2)}_{\nu} + \nabla^{(0)}_{\nu}\xi^{(2)}_{\mu} - g^{(0)}_{\mu\nu}\nabla^{(0)\lambda}\xi^{(2)}_{\lambda}) \\ &+ 2R^{(0)}_{\lambda\mu\sigma\nu}(\nabla^{(0)\lambda}\xi^{(2)\sigma} + \nabla^{(0)\sigma}\xi^{(2)\lambda}).\end{aligned} \tag{4.43}$$

Simplifying according to the commutation of the covariant derivatives, and imposing that the entire calculation be taken in vacuum, where $R^{(0)}_{\mu\nu} = 0$, we find

$$\begin{aligned}E^{(0)}_{\mu\nu}[\bar{h}^{(2)}] - E^{(0)}_{\mu\nu}[\bar{h}'^{(2)}] =& -\nabla_{\mu}\left(w_{\lambda}(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_A)\bar{h}^{(1)\lambda}_{\nu}\right) - \nabla_{\nu}\left(w_{\lambda}(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_A)\bar{h}^{(1)\lambda}_{\mu}\right) \\ &+ g_{\mu\nu}(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_A)\bar{h}^{(1)\lambda\rho}\nabla_{\rho}w_{\lambda}.\end{aligned} \tag{4.44}$$

This completes the demonstration that the two natural choices of multiscale promotion of a physical gauge condition, (4.37) and (4.38), are related by a second-order gauge transformation $\xi^{(2)}$ compatible with the multiscale construction.

4.3.6.2 Large, slow time dependent gauge transformations

In addition to the $\mathcal{O}(\varepsilon)$ gauge transformations $\xi(x^i, q^A, \tilde{w})$, which are a promotion of the typical perturbative gauge freedom to the multiscale domain, there is a separate class of large gauge

transformations in the multiscale formalism. Consider a large, global isometry of the background spacetime of the form $\Lambda^{\mu}{}_{\nu}(\tilde{w}), t^{\mu}(\tilde{w})$, corresponding to a slow time dependent rotations, boosts, and translations. We write the linear and angular momentum of the central Kerr black hole as P^{μ} and $J^{\mu\nu}$, which then transform under $\Lambda^{\mu}{}_{\nu}(\tilde{w}), t^{\mu}(\tilde{w})$ as:

$$P^{\mu} \rightarrow \Lambda^{\mu}{}_{\nu} P^{\nu}, \tag{4.45a}$$

$$J^{\mu\nu} \rightarrow \Lambda^{\mu}{}_{\sigma} \Lambda^{\nu}{}_{\rho} (J^{\sigma\rho} - 2t^{[\mu} P^{\nu]}). \tag{4.45b}$$

We take the restricted set of these transformations that preserves the form of the background Kerr metric, so consists only of time translations $t \rightarrow t + F_t(\tilde{w})$ and axial rotations $\phi \rightarrow \phi + F_{\phi}(\tilde{w})$.

We now consider the effect of these transformations on the metric perturbation, $h^{(1)}[q^A, \tilde{w}, r, \theta, \phi]$. The effect of the slow time dependent time translation is then,

$$\begin{aligned} q^A = \frac{1}{\varepsilon} \int^{\tilde{w}} \Omega^A \rightarrow \frac{1}{\varepsilon} \int^{\tilde{w}+\varepsilon F_t(\tilde{w})} \Omega^A, \\ \Rightarrow q^A \rightarrow q^A + F_t(\tilde{w}) \Omega^A(\tilde{w}). \end{aligned} \tag{4.46}$$

As the shift in fast time variable is constant at fixed $\tilde{w}$, the metric perturbation remains periodic in the new angle variables. The timelike components of the first order metric perturbation must also be corrected,

$$h^{(1)}_{\alpha\beta}[q^A, \tilde{w}, r, \theta, \phi] \rightarrow h^{(1)}_{\alpha\beta}[q^A + F_t(\tilde{w}) \Omega^A(\tilde{w}), \tilde{w}, r, \theta, \phi + F_{\phi}(\tilde{w})] + S_{\alpha\beta}(\tilde{w}) + \mathcal{O}(\varepsilon), \tag{4.47}$$

where

$$S = 2 g_{w\mu} F_t'(\tilde{w}) dw dx^{\mu} + 2 g_{\phi\mu} F_{\phi}' d\phi dx^{\mu} \tag{4.48}$$

Finally, we consider the effects of the inclusion of these large, slow transformations on a mode decomposition of the various physical quantities. As each physical variable is definitionally 2π periodic in q^A and in ϕ, we may define a mode decomposition for any physical quantity f:

$$f(x^i, q^A, \tilde{w}) = \sum_{k^A, m} A_{m, k^A}(\theta, r, \tilde{w}) e^{im\phi} e^{ik_A q^A}. \tag{4.49}$$

The remaining dependence on F_t and F_{ϕ} have the effect of multiplying each term in a mode expansion by a $\tilde{w}$-dependent quantity:

$$\sum_{k^A, m} A_{m, k^A}(\tilde{w}) e^{im\phi} e^{ik_A q^A} \rightarrow \sum_{k^A, m} \left[A_{m, k^A}(\tilde{w}) e^{im F_{\phi}(\tilde{w})} e^{ik_A \Omega^A F_t(\tilde{w})} \right] e^{im\phi} e^{ik_A q^A}, \tag{4.50}$$

which preserves the magnitude of the mode coefficients in either the Lorenz gauge expansion or the Teukolsky-Lousto-Campanelli expansion, but introduces a slow time dependent phase factor between the original and gauge-transformed expansions.

4.4 The self consistent approximation in the Body Zone

4.4.1 Self force theory in the self consistent formalism

In this section we review the self-consistent formalism of an EMRI computation [256], which we use to compute the acceleration of the small companion and the back-reaction to the metric perturbation. Deviations from the background Kerr metric $g^{(0)}_{\mu\nu}$ are characterized by the ratio of the mass of the small companion to the mass of the central black hole $\mu/M \equiv \varepsilon \ll 1$. The metric ansatz for an EMRI consists of an approximate metric perturbation $h_{\mu\nu}$ expressed as a functional of the worldline γ of the small companion:

$$g_{\mu\nu} \equiv g^{(0)}_{\mu\nu} + \varepsilon h_{\mu\nu} \equiv g^{(0)}_{\mu\nu} + \varepsilon h^{(1)}_{\mu\nu}[\gamma, x^\mu] + \varepsilon^2 h^{(2)}_{\mu\nu}[\gamma] + \mathcal{O}(\varepsilon^3). \tag{4.51}$$

Importantly, the expansion in ε is taken with the worldline γ held fixed.

The full Einstein field equation,

$$G_{\mu\nu}[h] = 8\pi T_{\mu\nu}, \tag{4.52}$$

is satisfied only when γ is the full, self-forced worldline. A gauge condition is imposed to reduce the full Einstein field equation to the relaxed Einstein field equation. The relaxed Einstein equation is solvable for an arbitrary source, and in particular for a small body following an arbitrary wordline. The gauge condition constrains the remaining equation of motion for the worldline itself. Consider a generic gauge condition:

$$L_\mu[h_{\mu\nu}] = 0. \tag{4.53}$$

The relaxed Einstein operator associated with the gauge condition is denoted:

$$G_{\mu\nu}|_{L_{\mu\nu}[h]=0} \equiv -2E_{\mu\nu}[h] \tag{4.54}$$

The metric perturbation $h_{\mu\nu}[\gamma]$ near the small companion can be expressed as the sum

$$h_{\mu\nu} = h^{\mathcal{P}}_{\mu\nu} + h^{\mathcal{R}}_{\mu\nu} = \varepsilon \left[h^{\mathcal{P}(1)}_{\mu\nu} + h^{\mathcal{R}(1)}_{\mu\nu}\right] + \varepsilon^2 \left[h^{\mathcal{P}(2)}_{\mu\nu} + h^{\mathcal{R}(2)}_{\mu\nu}\right] + \mathcal{O}(\varepsilon^3) \tag{4.55}$$

of a puncture metric $h^{\mathcal{P}}_{\mu\nu}$ and a residual metric $h^{\mathcal{R}}_{\mu\nu}$. The residual metric is a smooth, vacuum solution on γ, and is chosen such that it contains no negative powers of $\tilde{r}$ in a Fermi-Walker expansion about γ. The residual field so defined that was calculated in [250] gives rise to a self-forced worldline γ that is geodesic in the full vacuum solution $g^{(0)} + h^{\mathcal{R}}_{\mu\nu}$ [287].

The details of the puncture metric are determined by a matched asymptotic expansion of the metric near the small companion in powers of $\tilde{r}/\varepsilon$ and an expansion outside the small companion's puncture region in powers of $\tilde{r}$. The consistency of these expansions in an overlap region fully constrains the puncture metric up to a small number of parameters that describe the mass moments and spin of the small object. In this paper, we restrict attention to non-spinning spherically symmetric bodies.

For the purposes of the computation presented in this paper, we will solve the wave equations for the residual field using an effective source obtained from puncture metric [250]:

$$E_{\mu\nu}[h^{\mathcal{R}(1)}] = -16\pi \bar{T}^{(1)}_{\mu\nu}[\gamma] - E_{\mu\nu}[h^{\mathcal{P}(1)}] = 0 \qquad \text{Puncture zone} \tag{4.56a}$$

$$E_{\mu\nu}[h^{(1)}] = 0 \qquad \text{Interaction zone} \tag{4.56b}$$

$$E_{\mu\nu}[h^{\mathcal{R}(2)}] = 2\delta^2 R_{\mu\nu}[h^{\mathcal{R}(1)} + h^{\mathcal{P}(1)}, h^{\mathcal{R}(1)} + h^{\mathcal{P}(1)}] - 16\pi \bar{T}^{(2)}_{\mu\nu}[\gamma] - E_{\mu\nu}[h^{\mathcal{P}(2)}] \qquad \text{Puncture zone} \tag{4.56c}$$

$$E_{\mu\nu}[h^{(2)}] = 2\delta^2 R_{\mu\nu}[h^{(1)}, h^{(1)}] \qquad \text{Interaction zone} \tag{4.56d}$$

$$h^{(n)}_{\mu\nu} = h^{\mathcal{R}(n)}_{\mu\nu} + h^{\mathcal{P}(n)}_{\mu\nu} \qquad \text{Matching condition} \tag{4.56e}$$

Here the overbar denotes trace reversal, indices are raised and lowered using the background metric $g^{(0)}_{\mu\nu}$ and covariant derivatives are taken with respect to this metric.

As demonstrated in [287], the matching condition along with the gauge constraint and the form of the metric in the vicinity of the small companion can be used to derive the self force through second order:

$$\frac{D^2 z^\mu}{D\tau^2} = -\frac{1}{2}\left(g^{\mu\nu} - u^\mu u^\nu\right)\left(g^\gamma_\nu - h^{\mathcal{R}\gamma}_\nu\right)\left(2h^{\mathcal{R}}_{\gamma\alpha;\beta} - h^{\mathcal{R}}_{\alpha\beta;\gamma}\right)u^\alpha u^\beta. \tag{4.57}$$

Equations (4.56) comprise a set of coupled equations allowing us to solve self-consistently for both the worldline $z^\mu(\tau)$ and the first and second order metric perturbations. Equations (4.56a) and (4.56c) are solved using retarded boundary conditions and determine the self-force given by Eq.

(4.57) to second order. The first order contribution to the self force is derived from the field equations by evaluating the gauge condition (4.53) and taking advantage of the Bianchi identity [256]. The second order contribution to the self-force is derived by performing an inner and outer expansion of the metric near the small companion, and finding the gauge vector required to bring the pair of expansions into agreement [248, 249].

The metric perturbations $h^{(1)}_{\mu\nu}$ and $h^{(2)}_{\mu\nu}$ are obtained by solving the homogeneous field equations (4.56b) and (4.56d). In the common matching region, where the distance to the small companion is $\mu \ll \tilde{r} \ll M$, the matching condition (4.56e) is enforced. This matching resolves the remaining degrees of freedom associated with boundary conditions of the finite regions in which we solve the linearized, gauge-fixed Einstein field equations.

4.4.2 Limited time domain of validity for self consistent

The self-consistent formalism handles correctly the secular behavior of the worldline by expanding the metric perturbation at fixed worldline (4.51). With this technique, even at late times, the $\mathcal{O}(M)$ change in the worldline position will cause no breakdown of the formalism. Indeed, the worldline need not obey virtually any simplifying assumptions for the self-consistent formalism to remain a faithful approximation.

However, the self-consistent formalism makes the subtle assumption that the secular effects may be written entirely as functionals of the worldline. Extreme mass ratio inspirals violate this assumption via the slow evolution of the background spacetime. The central black hole will accrete mass and spin of order $\mathcal{O}(\mu)$ from the perturbations sourced over the full duration of the inspiral. The eventual magnitude of effect of the mass and spin shift is comparable to the $\mathcal{O}(\varepsilon)$ first order metric perturbation. The self-consistent formalism derives the flux of energy and spin (correctly) as a second order quantity, and therefore also derives (incorrectly, at late times) the mass and spin shift due to that flux also as a second order quantity.

A further complication to these slow shifts of the background in the self-consistent formalism is that they are not determined by the wave equation. At each order, the wave equations (4.56) are solvable for any perturbation of the central mass and spin, so offer no constraints on these secular effects. Instead, the time derivative of central mass and spin are fixed by the gauge condition (4.53).

As the relaxed Einstein field equations are constraint-preserving, the gauge is fixed for the initial data surface in the self-consistent computation, and hold for the full evolution of the system. This formalism grants no freedom to the evolution of the central mass, which grows linearly for the full duration of the system at the rate fixed by initial data.

By times $t \sim \mathcal{O}(M/\varepsilon)$, the solutions for the metric perturbation develops $\mathcal{O}(\varepsilon^2)$ errors associated with neglecting the contributions from the quadratic source

$$4\delta^2 R_{\mu\nu}[h^{(1)}, h^{\delta M} + h^{\delta a}] + \delta^2 R_{\mu\nu}[h^{\delta M} + h^{\delta a}, h^{\delta M} + h^{\delta a}]. \tag{4.58}$$

To describe the inspiral reliably at these late times, the self-consistent formalism must be modified to account for the slowly varying mass and spin of the central black hole. We may either accomplish this by augmenting the self consistent procedure with elements of the multiscale computation, or by performing an expansion of the multiscale results for small intervals in slow time and matching to the self consistent solution at similar times. Both methods are discussed in more detail in the appendix (4.C)

4.4.3 | Coordinate time action and angle variable formalism for Kerr geodesics

We build the multiscale orbital motion as slow variations of and small perturbations from the Kerr geodesic. In the limit of 'turning off' the gravitational backreaction (taking $\tilde{w}$ to be constant), the leading order orbital motion is geodesic. The action and angle form for the conservative Kerr geodesic is a convenient starting point for the orbital computation. The action and angle construction is presented in more detail in [222, 299] , but for completeness we review here the formalism required for the two-timescale orbital computation.

Define phase space coordinates (q_α, p_α) and a Hamiltonian H. For geodesics in Kerr, there are three first integrals of motion: energy E, angular momentum L_z, and the Carter Constant Q, each of which is conserved for geodesic motion. We will write these first integrals, along with the Hamiltonian H as $P_\alpha = (H, E, L_z, Q)$.

An implication of the more general derivation by [300] is that the phase space of the Kerr geodesic is foliated by the level sets of constant P_α. Further, these level sets are diffeomorphic to $T^3 \times \mathbb{R}$, where the only non-compact coordinate direction describes the time coordinate of the geodesic. The

coordinates on the phase space can be chosen to be the generalized action-angle variables (q_α, J_α). As the range of the time component of the coordinates q_α is non-compact over the geodesic motion, we denote a 3+1 separation of these variables as $q_\alpha = (q_t, q_i)$. The action and angle variables may also be chosen such that each q_i is 2π periodic, and the first integrals of motion depend exclusively on the action variables $P_\alpha = P_\alpha(J_\beta)$.

For this choice of generalized action-angle variables, the equations of motion are:

$$\begin{aligned} \dot{q} =& \frac{\partial H(J)}{\partial J_\alpha} \\ \dot{J}_\alpha =& -\frac{\partial H(J)}{\partial q_\alpha} = 0 \end{aligned} \tag{4.59}$$

For each angular variable, we define the quantity

$$\Omega_\alpha(J) \equiv \frac{\partial H(J)}{\partial J_\alpha}, \tag{4.60}$$

The set of action and angle variables q_α, J_β parameterize an 8-dimensional phase space. However, the momentum of the small companion is constrained by the normalization condition $g^{\mu\nu}p_\mu p_\nu = \mu^2$, which relates the magnitude of the momentum four-vector to one of the constants of motion. The dynamics we wish to derive can be written as the evolution of three momentum degrees of freedom. The calculations of the metric and waveforms in the next sections will require a global time parameter rather than the proper time of the worldline.

We therefore re-parameterize the dynamics by performing a canonical phase space reduction to a 6−dimensional submanifold, which remains symplectic. This restricts the motion to the surface $H = -\mu^2/2$, where H is the Hamiltonian for geodesic dynamics in Kerr, and leads to three angle variables q_i together with the conserved quantities J_i. We can take advantage of this phase space reduction to also write the dynamics in terms of the modified Boyer-Lindquist coordinate time (4.155) $\{p_i(w), x^i(w)\}$. The new Hamiltonian can be written as:

$$H = \frac{g^{wi}}{g^{ww}}p_i + \frac{1}{|g^{ww}|}\left[(g^{wi}p_i)^2 - g^{ww}(g^{ij}p_ip_j + \mu^2)\right]^{1/2} \tag{4.61}$$

The set of first integrals for the new Hamiltonian system is $P_i = \{E, L_z, K\}$, where E is now the new Hamiltonian, and μ^2 is treated as a non-dynamical constant. The action variables are defined from the symplectic potential of the phase space associated with the new Hamiltonian. The

resulting set of three action variables is then

$$\begin{aligned} J_r =& \frac{1}{2\pi}\oint E\,h'(r) + \frac{\sqrt{\hat{V}_r}}{\Delta}dr \\ J_\theta =& \frac{1}{2\pi}\oint \sqrt{\hat{V}_\theta}d\theta \\ J_\phi =& L_z, \end{aligned} \tag{4.62}$$

where the potentials $\hat{V}_r, V_\theta$ are

$$\begin{aligned} \hat{V}_r &= (\varpi^2 E - aL_z)^2 - \Delta(\mu^2 r^2 + K) \\ \hat{V}_\theta &= K - \mu^2 a^2 \cos^2\theta(L_z - aE\sin^2\theta). \end{aligned} \tag{4.63}$$

The three relations (4.62) give a one-to-one map between the action variables J_i and the constants of motion P_i. In principal, these relations are invertible to find

$$P_i = P_i(J_j). \tag{4.64}$$

The angle variables are obtained from a canonical transformation with a generating function

$$\mathcal{W} = L_z\phi + q_r\mathcal{W}_r(r) + q_\theta\mathcal{W}(\theta), \tag{4.65}$$

where

$$\begin{aligned} \mathcal{W}_r(r) &= \int^r \frac{\sqrt{\hat{V}_r}}{\Delta}dr \\ \mathcal{W}_\theta(\theta) &= \int^\theta \sqrt{\hat{V}_\theta}d\theta. \end{aligned} \tag{4.66}$$

The canonical transformation so defined takes the coordinates $\{x^i, p_i\}$ to $\{q^i, J_i\}$, with above defined J_i (4.62) and

$$q_i = \frac{\partial\mathcal{W}}{\partial J_i} \tag{4.67}$$

Due to the one-to-one map between the action variables J_i and the first integrals of motion P_i, we will choose to express the dynamics of the system by equations for the time derivatives of $\{q_i, P_i\}$. The geodesic equations of motion are then

$$\frac{dq_i}{dt} = \frac{dE}{dJ_i} \equiv \omega_i(P_i) \tag{4.68a}$$

$$\frac{dP_i}{dt} = 0 \tag{4.68b}$$

In much of the notation which follows we optionally denote the components of the action-angle variables with capital roman characters, as q^A, J^M or P^M, which is used to indicate more generally either the set of three components, or fewer if the orbit or background is appropriate, and to distinguish the index from the spatial components of spacetime vectors which will also use lowercase roman characters.

4.4.4 A summary of prior results: multiscale orbital evolution

We briefly summarize the multiscale analysis of the orbital motion presented in [222]. The first step uses the Hamiltonian structure of the unperturbed, geodesic motion to rewrite the forced equations of motion in terms of generalized action angle variables. The equation of motion for the self-forced worldline is

$$\frac{d^2x^\nu}{d\tau^2} + \Gamma^\nu_{\sigma\rho}\frac{dx^\sigma}{d\tau}\frac{dx^\rho}{d\tau} = \varepsilon a^{(1)\,\nu} + \varepsilon^2 a^{(2)\,\nu} + O(\varepsilon^3), \tag{4.69}$$

where τ is the worldline proper time, and $a^{(1)\,\nu}$ and $a^{(2)\,\nu}$ are the first order and second order self-accelerations obtained from the use of the self force equations [248, 249, 254, 255] with the metric perturbation obtained from the multiscale computation (4.56). Imposing the relation between $\{x^i, p_j\}$ and $\{q^A, P^M\}$ expressed in subsection A, the action and angle equations of motion for a self-forced worldline are

$$\frac{dq^A}{dt} = \omega^A(P^M) + \varepsilon g^{A(1)}(q^r, q^\theta, P^M) + \varepsilon^2 g^{A(2)}(q^r, q^\theta, P^M) + O(\varepsilon^3), \tag{4.70a}$$

$$\frac{dP^M}{dt} = \varepsilon G^{M(1)}(q^r, q^\theta, P^M) + \varepsilon^2 G^{M(2)}(q^r, q^\theta, P^M) + O(\varepsilon^3). \tag{4.70b}$$

Here the variables P_i are the three parameters of geodesic motion. In the self-force context, they are no longer conserved, but are taken to have the same functional relationship to the forced $\{x, p\}$ as their conserved counterparts had to the geodesic $\{x, p\}$:

$$E(q^A, \tilde{w}) \equiv p_\mu \xi^\mu_w \tag{4.71a}$$

$$L_z(q^A, \tilde{w}) \equiv p_\mu \xi^\mu_\phi \tag{4.71b}$$

$$K(q^A, \tilde{w}) \equiv K^{\mu\nu} p_\mu p_\nu, \tag{4.71c}$$

where ξ^μ_t and ξ^μ_ϕ are the timelike and axial killing vectors, respectively, and $K^{\mu\nu}$ is the Kerr killing tensor.

The functions $g^{A(1)}$, $G^{M(1)}$ are determined by the first order self acceleration, and similarly, $g^{A(2)}$ and $G^{M(2)}$ are determined in part by the second order self acceleration.

The forcing terms for the angle variables are determined by computing the required forcing terms in the (x, p) Hamilton-Jacobi equations of motion, then performing the canonical transformation to (q, J). The computation necessary for determining g^A is detailed in in appendix A. We perform a similar computation to recover the forcing terms G^M. However, due to our choice to use P^M directly, there is a simpler method for determining the forcing expressions. For the energy and angular momentum, we can write:

$$\frac{dE}{dw} = \frac{d}{dw}(u_\alpha \xi^\alpha_w) = \frac{d\tau}{dw}\frac{d}{d\tau}(u_\alpha \xi^\alpha_w) = \frac{d\tau}{dw} a_\alpha \xi^\alpha_w = \left(-g_{ww} - 2g_{wi}v^i - g_{ij}v^i v^j\right)^{1/2} a_w \tag{4.72a}$$

$$\frac{dL_z}{dw} = \frac{d}{dw}(u_\alpha \xi^\alpha_\phi) = \frac{d\tau}{dw}\frac{d}{d\tau}(u_\alpha \xi^\alpha_\phi) = \frac{d\tau}{dw} a_\alpha \xi^\alpha_\phi = \left(-g_{ww} - 2g_{wi}v^i - g_{ij}v^i v^j\right)^{1/2} a_\phi, \tag{4.72b}$$

where u is the normalized worldline four-velocity $u^\alpha = Dz^\alpha/d\tau$ and v^i is the three-velocity with respect to the time coordinate w, $v^i = dz^i/dw$. The forcing term derivation for the Carter constant may also be simplified, using instead the Killing tensor:

$$\begin{aligned}\frac{dK}{dw} &= \frac{d\tau}{dw}\frac{d}{d\tau}\left(2K_{\alpha\beta}u^\alpha u^\beta\right) = \frac{d\tau}{dw} K_{\alpha\beta} a^\alpha u^\beta \\ &= (-g_{ww} - 2g_{wi}v^i - g_{ij}v^i v^j)^{1/2}\left(K_{\alpha w}a^\alpha + K_{\alpha i}a^\alpha v^i\right)\end{aligned} \tag{4.73}$$

We next apply the method of multiscale expansion described in the book [293], and summarized in Section IIB. A multiscale expansion procedure is an effective systematic method for studying the cumulative effect of a small disturbance active over a long time such as the effect of a small damping on an oscillator. First, as is justified in [222], we directly associate the three fast time variables of the system φ^A with the action-angle variables $\varphi^A = q^A$.

The ansatz for q^A and J^M as a multiscale expansion expresses the dependence on (w, ε) as an expansion in ε with the fast dynamics of order off the expansion represented by the phases q^A and the slow evolution represented by the dependence on $\tilde{w}$. Specifically, we assume expansion in $\tilde{w}$,

$$P^{M(n)}(q^A, \tilde{w}) = P^{(0)M}(q^A, \tilde{w}) + \varepsilon P^{(1)M}(q^A, \tilde{w}) + O(\varepsilon^2). \tag{4.74}$$

The expansion coefficients $P_i^{M(n)}$ are multiply periodic in the phase variables q^A with period 2π in each variable so that $P^{M(n)}(q^A + 2\pi k^A, \tilde{w}) = P_i^{(n)}(q_j, \tilde{w})$, for any k^A N-tuple of integers.

We substitute the ansatz (4.74) into the equations of motion (4.70) and solve order by order in ε. At each order we decompose the governing equations into an averaged and an oscillatory piece which we solve separately and obtain unique solutions.

To discuss the expansion of the results for the adiabatic order computation, it is convenient to introduce the fast time averaging operation. The average of any function on phase space, $F(q^A, J^M)$ is defined as

$$\langle F \rangle \equiv \frac{1}{(2\pi)^2} \int_0^{2\pi} dq^r \int_0^{2\pi} dq^\theta \, F(q^r, q^\theta, J^M). \tag{4.75}$$

The multiscale prescription determines the expansion coefficients in terms of the forcing functions. At adiabatic order, the functions $qA(0)$ and $\mathcal{P}^{M(0)} \equiv \langle P^{M(0)} \rangle$ are obtained from the leading set of equations,

$$\frac{dq^{A(0)}}{d\tilde{w}} = \omega^A[\mathcal{P}^{M(0)}(\tilde{w})], \tag{4.76a}$$

$$\frac{d\mathcal{P}^{M(0)}}{d\tilde{w}} = \left\langle G^{M(1)} \right\rangle [\mathcal{P}^{M(0)}(\tilde{w})]. \tag{4.76b}$$

These equations imply that at adiabatic order, for each fixed $\tilde{w}$, the orbit is precisely geodesic in q^A, with slowly varying orbit parameters $P^M(\tilde{w})$. Finally, the slow variation of those orbit parameters can, in turn, be determined by asymptotic fluxes, as they depend only on the dissipative part of the first order self force.

The subleading, post-adiabatic corrections to the inspiral are determined from differential equations that depend on (i) all pieces of the first order self force, and (ii) the fast-time averaged, dissipative piece of the second order self force. The full expansion of the orbit was previously presented in [222], and a discussion of the full Interaction Zone picture to post-adiabatic order is presented in Chapter 5.

4.5 Geometric optics expansion in the far zone

4.5.1 Preamble

Far from the central black hole and the inspiraling companion ($M \ll r$), the multiscale ansatz (4.15) is no longer valid. The assumption made to construct that ansatz is that there is no scale longer than the radiation-reaction timescale, which is violated by the $\propto 1/r$ dependence at $r \sim M/\varepsilon$. Instead,

we adopt a geometric optics formalism, which is well-suited to the new separation of lengthscales between the $\sim M$ wavelength of the outgoing radiation and the comparatively long scale of the remaining relevant scales in the Far Zone:

- The local radius of curvature of spacetime $\sim \sqrt{r^3/M}$;
- The radius of curvature $\sim r$ of wavefronts and the lengthscale $\sim r$ over which the amplitude of the radiation evolves;
- The timescale $\tilde{w} \sim M/\varepsilon$ over which the frequencies of radiation evolve;

In this section we will introduce the formalism for the geometric optics approximation, which is a minor extension to the standard textbook treatment [289], and the results at the order necessary to derive adiabatic-order effects. The key result is that, at adiabatic order, a direct treatment of geometric optics propagation on the linearized, monopolar expansion of the Kerr metric is sufficiently accurate to describe the spacetime all the way to future null infinity, and directly matches onto an Interaction Zone Lorenz gauge solution.

4.5.2 | Ansatz for the metric

We assume a one parameter family $g_{\alpha\beta}(\varepsilon)$ of vacuum solutions labeled by the mass ratio $\varepsilon = \mu/M$ that satisfies assumptions 1 and 2 of Sec. 4.2.1. We assume that in the Far Zone there exists a smooth function $\Theta(\tilde{x}^\gamma)$, and a class of coordinate systems $\tilde{x}^\alpha$, for which the metric takes the form

$$g_{\alpha\beta}(\tilde{x}^\gamma, \varepsilon) = \varepsilon^{-2}\mathcal{G}_{\alpha\beta}(\tilde{x}^\gamma, \varpi, \varepsilon), \tag{4.77}$$

where the function $\mathcal{G}_{\alpha\beta}$ is smooth in all of its arguments, periodic in its second argument ϖ with period 2π, and

$$\varpi = \frac{\Theta(\tilde{x}^\gamma)}{\varepsilon}. \tag{4.78}$$

We assume that the function $\mathcal{G}$ has an asymptotic expansion in powers of ε near $\varepsilon = 0$:

$$\mathcal{G}_{\alpha\beta} = \eta_{\alpha\beta} + \varepsilon h_{\alpha\beta} + \varepsilon^2 j_{\alpha\beta} + \varepsilon^3 k_{\alpha\beta} + \varepsilon^4 l_{\alpha\beta} + O(\varepsilon^5), \tag{4.79}$$

where all quantities are functions of $\tilde{x}^\gamma$ and ϖ, and the $O(\varepsilon^4)$ refers to the limit $\varepsilon \to 0$ at fixed $\tilde{x}^\alpha$ and fixed ϖ. Finally we assume that $\eta_{\alpha\beta}$ is flat and that $\eta_{\alpha\beta}$ and $h_{\alpha\beta}$ are independent of ϖ.

4.5.3 Periodic decomposition

Since the metric perturbations $j_{\alpha\beta}$ and $k_{\alpha\beta}$ are periodic functions of the phase variable ϖ, they can be decomposed in a harmonic series in the usual way:

$$j_{\alpha\beta}(x^\gamma, \varpi) = \sum_{n=-\infty}^{\infty} j_{n\,\alpha\beta}(x^\gamma) e^{in\varpi}, \tag{4.80}$$

with

$$j_{n\,\alpha\beta}(x^\gamma) = \frac{1}{2\pi}\int_0^{2\pi} d\varpi e^{-in\varpi} j_{\alpha\beta}(x^\gamma, \varpi). \tag{4.81}$$

We write this decomposition in the form

$$j = j_0 + \delta j. \tag{4.82}$$

Here j_0 is the $n = 0$ terms which are independent of ϖ, and δj contains all the oscillatory terms with $n \neq 0$. We also define an averaging operation $\langle \ldots \rangle$ on tensor fields which are functions of $\tilde{x}^\alpha$ and ϖ via

$$\langle f \rangle \equiv \frac{1}{2\pi}\int_0^{2\pi} d\varpi f(\varpi). \tag{4.83}$$

We then have

$$\langle j_{\alpha\beta} \rangle = j_{0\,\alpha\beta}, \tag{4.84}$$

In the following discussion, we denote derivatives with respect to the phase variable Θ as

$$\partial_{(\Theta/\varepsilon)} f(\ldots, \Theta/\varepsilon) = f'(\ldots, \Theta/\varepsilon). \tag{4.85}$$

In particular, a derivative with respect to scaled coordinates becomes,

$$\partial_\mu (f(\tilde{x}^\mu, \Theta/\varepsilon)) = \bar{\partial}_\alpha f(\tilde{x}^\mu, \Theta/\varepsilon) + \frac{1}{\varepsilon} l_\mu f'(\tilde{x}^\mu, \Theta/\varepsilon), \tag{4.86}$$

where the derivative $\bar{\partial}_\alpha$ is evaluated at fixed Θ and l_α is the wavevector,

$$l_\alpha(\tilde{x}^\mu) = \nabla_\alpha \Theta(\tilde{x}^\mu) \tag{4.87}$$

4.5.4 | Discussion

In the Far Zone, we adopt scaled coordinates $\tilde{x}^\mu = \varepsilon x^\mu$, which encode the long spatial scales relevant to the metric perturbations and the long timescales relevant to the radiation-reaction evolution of the outgoing modes. Meanwhile, the dependence on the phase variable Θ encodes the rapid variation of the radiation on a lengthscale $\sim M$. The first term $\eta_{\alpha\beta}$ in the expansion (4.79) will simply be the Minkowski metric, and the second term $h_{\alpha\beta}$ will be the leading expansion of the background Kerr metric at large $\tilde{r}$. The radiation first arises in the third term in the expansion $j_{\alpha\beta}$, due to the leading $\sim 1/\tilde{r}$ dependence of the radiation leading to a promotion as compared to the $\sim \varepsilon$ values in the Interaction Zone .

The expansion (4.79) is nominally valid for $\tilde{r} \sim \varepsilon^0$. However we will show below that it is valid for smaller $\tilde{r}$, all the way down to

$$\tilde{r} \gg \varepsilon M, \quad or \quad r \gg M, \tag{4.88}$$

the inner edge of the far zone. In addition we will show in Sec. 4.5.9 below that a minor extension of the approximation has a domain of validity that extends all the way out to $r \to \infty$.

We note that the ansatz (4.79) is essentially the textbook ansatz for the geometric optics approximation given in, for example, Misner, Thorne and Wheeler [289], except for a number of differences:

- The leading oscillatory term has a specific scaling $\propto \varepsilon^2$, where ε is the typical wavelength of the oscillations.
- There is the non-oscillatory term $h_{\alpha\beta}$ which enters at an order intermediate between the background and the oscillatory term.
- There is a non-oscillatory part of the perturbations j_0, which is required for the post-adiabatic computation. Its derivation is performed by techniques closely related to [301] and [278], and is presented in Chapter 6.
- The textbook treatments usually keep only the $n = 1$ term in the harmonic decomposition. The higher order terms will be needed here in order to capture the full gravitational waveform.

4.5.5 Gauge freedom and specialization

General gauge transformations consist of diffeomorphisms

$$\chi_\varepsilon : \mathcal{M} \to \mathcal{M}, \tag{4.89}$$

where $\mathcal{M}$ is the spacetime manifold, which depend on ε. The requirement that the metric satisfy our geometric optics ansatz (4.77) and (4.79) restricts the dependence on ε of the allowed gauge transformations. In this section we discuss the gauge freedom and we also make some further gauge specializations that will simplify the form of the Einstein equations.

We assume that the gauge transformation has a well defined limit as $\varepsilon \to 0$:

$$\chi_0 = \lim_{\varepsilon \to 0} \chi_\varepsilon. \tag{4.90}$$

It is possible that more general gauge transformations that preserve our ansatz do exist, but we will not consider them here. The diffeomorphism χ_0 acts on the Minkowski background metric $\eta_{\alpha\beta}$ and can be used to express it in any coordinate system. We assume that we have fixed a background coordinate system and restrict attention from now on to transformations where χ_0 is the identity map. We further specialize to the set of gauge transformations which also preserve the first order $h_{\alpha\beta}$ determined by expansion of the background Kerr metric.

We can then express χ_ε in terms of a set of vector fields ζ^α, ω^α, $\chi^\alpha \ldots$, which are functions of x^α and $\varpi = \Theta/\varepsilon$, as

$$\chi_\varepsilon = \mathcal{D}_\zeta(\varepsilon^2) \circ \mathcal{D}_\omega(\varepsilon^3) \circ \mathcal{D}_\chi(\varepsilon^4) \circ \left[\mathbb{1} + O(\varepsilon^5)\right] \tag{4.91}$$

Here $\mathcal{D}_\xi(\varepsilon) : M \to M$ denotes the one parameter group of diffeomorphisms generated by ξ^α, which moves any point ε units along an integral curve of the vector field. In order to preserve the form of the metric ansatz (4.79), in which $\eta_{\alpha\beta}$ and $h_{\alpha\beta}$ are independent of ϖ, it will be necessary for the vector field ζ^α to be independent of ϖ.

We next consider the effects of the vector fields ζ^α and ω^α. We note that the Lie derivative of the background Minkowski metric with respect to ω^α can be written in the form

$$\mathcal{L}_\omega \eta_{\alpha\beta} = 2\bar{\nabla}_{(\alpha}\omega_{\beta)} + \frac{2}{\varepsilon} l_{(\alpha}\omega'_{\beta)}. \tag{4.92}$$

Using the expressions (4.91) and (4.92) we find that the field $j_{\alpha\beta}$ transforms as

$$j_{\alpha\beta} \to j_{\alpha\beta} + 2\bar{\nabla}_{(\alpha}\zeta_{\beta)} + 2l_{(\alpha}\omega'_{\beta)}. \tag{4.93}$$

We can rewrite this as separate transformation rules for j_0 and δj:

$$j_{0\,\alpha\beta} \to j_{0\,\alpha\beta} + 2\bar{\nabla}_{(\alpha}\zeta_{\beta)}, \tag{4.94}$$

and

$$\delta j_{\alpha\beta} \to \delta j_{\alpha\beta} + 2l_{(\alpha}\omega'_{\beta)}. \tag{4.95}$$

We can now use the vector field ζ^α to place the metric perturbation j_0 in any convenient gauge for the background. Separately, we can use the $n \neq 0$ components of ω^α to adjust the gauge of δj. The available gauge choices for δj, however, depend on the wavevector l_α. Therefore, we choose to leave the gauge condition unfixed until we have developed some of the details of the field equations in Sections 4.5.7 and 4.5.8. We choose to place j_0 in the Lorenz gauge, although the matching details and some of the remaining gauge degrees of freedom need not be fully determined until post-adiabatic order, which we discuss in Chapter 6.

4.5.6 | Expansion of Einstein equations

In the next section, we expand the full set of Einstein field equations in the multiscale geometric optics formalism. To describe the expansion of nonlinear differential operators, we introduce a notation to distinguish the order of expansion due to explicit factors of ε from the functional dependence on ε, and the perturbative expansion due to nonlinear products of subleading metric perturbations. We first denote the expansion in powers of the perturbation by,

$$G_{\mu\nu}[g^{(0)} + \varepsilon h] = G_{\mu\nu} + \varepsilon\delta G_{\mu\nu}[h] + \varepsilon^2\delta^2 G_{\mu\nu}[h,h] + \varepsilon^3\delta^3 G_{\mu\nu}[h,h,h] + \mathcal{O}(\varepsilon^4). \tag{4.96}$$

We denote with a superscript the number of powers of ε which arise from an explicit expansion of the covariant derivatives involving the phase Θ. Therefore, each order of the subleading differential operators $\delta^n G$, $n \geq 1$, we expand as

$$\delta^n G_{\mu\nu}[h\ldots] = \delta^n G^{(0)}_{\mu\nu}[h\ldots] + \frac{1}{\varepsilon}\delta^n G^{(-1)}_{\mu\nu}[h\ldots] + \frac{1}{\varepsilon^2}G^{(-2)}_{\mu\nu}[h\ldots], \tag{4.97}$$

where the superscript is chosen to be indicative of the number of powers of ε which should be included as a coefficient of the term. Note here that the maximum negative power of this is (-2), as we obtain one negative power of ε for each derivative ∂_Θ, and there are at most two derivatives in any term of the Einstein operator G. In the following subsection, we expand the Einstein field equation according to perturbative geometric optics framework, to the order necessary for a post-adiabatic analysis of the region $r \gg M$.

We now insert the metric ansatz given by (4.77) and (4.79) into the vacuum Einstein equations and expand in powers of ε, keeping terms through $O(\varepsilon^4)$. This yields (dropping indices for simplicity)

$$\begin{aligned} 0 = {} & \varepsilon\delta G[h] + \varepsilon^2\delta G[j] + \varepsilon^3\delta G[k] + \varepsilon^4\delta G[l] \\ & + \varepsilon^2\delta^2 G[h,h] + 2\varepsilon^3\delta^2 G[h,j] + 2\varepsilon^4\delta^2 G[h,k] \\ & + \varepsilon^4\delta^2 G[j,j] + \varepsilon^3\delta^3 G[h,h,h] + 3\varepsilon^4\delta^3 G^{(3)}[h,h,j] \\ & + \varepsilon^4\delta^4 G[h,h,h,h] + O(\varepsilon^5), \end{aligned} \tag{4.98}$$

where the background metric is $g^{\mathrm{B}}_{\alpha\beta} = \eta_{\alpha\beta}$. We next use the identity (4.97) for the operators $\delta^n G$, and re-expand in powers of ε, keeping terms up through $O(\varepsilon^2)$. This gives

$$\begin{aligned} 0 = {} & \delta G^{(-2)}[j] + \varepsilon\Big\{\delta G[h] + \delta G^{(-1)}[j] + \delta G^{(-2)}[k] \\ & + 2\delta^2 G^{(-2)}[h,j]\Big\} + \varepsilon^2\Big\{\delta G^{(0)}[j] + \delta G^{(-1)}[k] + \delta G^{(-2)}[l] \\ & + 2\delta^2 G^{(-1)}[h,j] + 2\delta^2 G^{(-2)}[h,k] + \delta^2 G[h,h] \\ & + \delta^2 G^{(-2)}[j,j] + 3\delta^3 G^{(-2)}[h,h,j]\Big\} + O(\varepsilon^3). \end{aligned} \tag{4.99}$$

Finally equating to zero the coefficients of the various powers of ε gives the following set of equations:

$$\delta G^{(-2)}[j] = 0, \tag{4.100}$$

$$\delta G[h] + \delta G^{(-1)}[j] + \delta G^{(-2)}[k] + 2\delta^2 G^{(-2)}[h,j] = 0. \tag{4.101}$$

and

$$\begin{aligned} 0 = {} & \delta G^{(0)}[j] + \delta G^{(-1)}[k] + \delta G^{(-2)}[l] + 2\delta^2 G^{(-1)}[h,j] \\ & + 2\delta^2 G^{(-2)}[h,k] + \delta^2 G[h,h] + \delta^2 G^{(-2)}[j,j] + 3\delta^3 G^{(-2)}[h,h,j]. \end{aligned} \tag{4.102}$$

We now analyze these equations one by one.

4.5.7 Leading order Einstein equation: null geodesic congruence

The leading term in the expansion (6.15) that is not explicitly fixed by a perturbatively expanded Kerr background in the Far Zone is δj. The δj contribution is fixed by the wave equation with two derivatives of the phase Θ,

$$\delta G^{(-2)}_{\mu\nu}[j] = 0. \tag{4.103}$$

We use the leading equation (4.103) field equation to argue that the vector l^μ must be null.

Seeking contradiction, assume first that the vector l^μ is not null. In this case, we may enforce the Lorenz gauge condition $l^\mu \delta j_{\mu\nu} = 0$. In the Lorenz gauge, the leading contribution to the Einstein field equation takes the convenient form,

$$(\eta^{\mu\nu} l_\mu l_\nu)\delta j''_{\lambda\sigma} = 0. \tag{4.104}$$

We match to solutions in the Interaction Zone which have outgoing wave modes, so the solution $\delta j_{\lambda\sigma} = 0$ does not satisfy the desired boundary conditions. Therefore, we may conclude that, for all nontrivial solutions in which we are interested, the wave vector is null $l_\mu l^\mu = 0$. The conclusion then contradicts the construction of a non-null l^μ, which implies that l^μ must always be null.

As we have established that the wave vector l is null, we may immediately define a null congruence associated with the null vector $\{l, n, m, \bar{m}\}$. We define these additional null vectors to be consistent with the Newman-Penrose formalism, reviewed in detail in appendix 5.C. In particular, we enforce,

$$n^\alpha n_\alpha = l^\alpha l_\alpha = m^\alpha m_\alpha = \bar{m}^\alpha \bar{m}_\alpha = 0, \tag{4.105a}$$

$$n^\alpha l_\alpha = -1 \qquad m^\alpha \bar{m}_\alpha = 1, \tag{4.105b}$$

and all other inner products of null vectors vanish. The various null vector projected components of the first covariant derivatives are given by the twelve spin coefficients $\{\alpha, \beta, \gamma, \varepsilon, \kappa, \lambda, \mu, \nu, \tau, \pi, \sigma, \rho\}$, also reviewed in detail also in appendix 5.C. In addition, it is useful to define the scalar coefficients for each tensor field,

$$q_{\alpha\beta} = q_{ll} n_\alpha n_\beta + q_{nn} l_\alpha l_\beta + 2q_{nl} n_{(\alpha} l_{\beta)} + q_{mm} \bar{m}_\alpha \bar{m}_\beta + q_{\bar{m}\bar{m}} m_\alpha m_\beta + 2q_{m\bar{m}} m_{(\alpha} \bar{m}_{\beta)} \tag{4.106}$$

$$- 2q_{nm} \bar{m}_{(\alpha} l_{\beta)} - 2q_{lm} \bar{m}_{(\alpha} n_{\beta)} - 2q_{n\bar{m}} m_{(\alpha} l_{\beta)} - 2q_{l\bar{m}} m_{(\alpha} n_{\beta)} \tag{4.107}$$

We choose to use the null tetrad in flat spacetime analogous to the Kinnersly tetrad of a Kerr background,

$$\vec{l} = \partial_t + \partial_r \tag{4.108a}$$

$$\vec{n} = \frac{1}{2}\partial_t - \frac{1}{2}\partial_r \tag{4.108b}$$

$$\vec{m} = \frac{1}{\sqrt{2}r}\partial_\theta + \frac{i}{\sqrt{2}\sin(\theta)r}\partial_\phi \tag{4.108c}$$

$$\vec{\bar{m}} = \frac{1}{\sqrt{2}r}\partial_\theta - \frac{i}{\sqrt{2}\sin(\theta)r}\partial_\phi \tag{4.108d}$$

which we use to explicitly evaluate all of the Newman-Penrose spin coefficients wherever they appear in intermediate steps of the geometric optics computation.

In the case of a null wavevector, we may not enforce the Lorenz gauge condition for the rapidly varying component of the metric perturbation δj. In this case, we take the weaker gauge condition $l^\alpha n^\beta \delta j_{\alpha\beta} = 0$, which may be used due to the form of the gauge transformation of the rapidly varying perturbations (4.95). In this gauge (or, indeed even without the additional gauge specification), the leading order Einstein field equation gives,

$$\delta G^{(-2)}_{nn}[j] = -\frac{1}{2}\delta j''_{m\bar{m}} = 0, \tag{4.109a}$$

$$\delta G^{(-2)}_{mn}[j] = -\frac{1}{2}\delta j''_{ml} = 0, \tag{4.109b}$$

$$\delta G^{(-2)}_{\bar{m}n}[j] = -\frac{1}{2}\delta j''_{\bar{m}l} = 0, \tag{4.109c}$$

$$\delta G^{(-2)}_{m\bar{m}}[j] = -\delta j''_{ll} = 0, \tag{4.109d}$$

$$\delta G^{(-2)}_{\mu l}[j] = 0, \tag{4.109e}$$

$$\delta G^{(-2)}_{mm}[j] = 0, \tag{4.109f}$$

$$\delta G^{(-2)}_{\bar{m}\bar{m}}[j] = 0, \tag{4.109g}$$

The leading wave equation then grants constraints on the components of δj,

$$\delta j_{\bar{m}m} = \delta j_{ll} = \delta j_{lm} = \delta j_{l\bar{m}} = 0. \tag{4.110}$$

Therefore, we learn from the leading order Einstein field equation that the wave vector l^μ is null, and that the oscillatory part of the metric perturbation δj obeys the Lorenz gauge condition $l^\mu \delta j_{\mu\nu}$; i.e. the perturbations are transverse, traceless, and propagate along outgoing null cones. As a final

note, there remains gauge freedom in the components of δj. However, we use this gauge freedom to significantly simplify the computation at the next order, at which the remaining components directly appear.

4.5.8 | Subleading order Einstein equation: propagation of radiation

At first subleading order, the expansion of the Einstein field equation expands to,

$$\delta G^{(-2)}_{\alpha\beta}[k] + \delta G_{\alpha\beta}[h] + \delta G^{(-1)}_{\alpha\beta}[j] + 2\delta^2 G^{(-2)}_{\alpha\beta}[h,j] = 0. \tag{4.111}$$

This form is considerably more complicated than the corresponding leading order form (4.103), and we will need to carefully extract a few different contributions to determine the full description of the outgoing waves.

First, consider the average of the above formula over 2π in the rapidly varying phase Θ/ε. This operation will remove all terms which contain only single terms with derivatives with respect to Θ, as each of those terms is necessarily of nonzero frequency. The result is simply,

$$\langle \delta G_{\alpha\beta}[h] \rangle = \delta G_{\alpha\beta}[h] = 0. \tag{4.112}$$

This equation ensures that h must be a solution of perturbing from a Minkowski background, which will certainly be satisfied if we simply construct h as the anticipated expansion of the Kerr metric in powers of M.

The remaining components can be split up into their respective Newman-Penrose components, much as in the leading expansion (4.109). Some of these give direct agreement with the Einstein field equations in the desired gauge, and so provide no additional information,

$$l^\mu \left(\delta G^{(-2)}_{\mu\beta}[k] + \delta G^{(-1)}_{\mu\beta}[j] + 2\delta^2 G^{(-2)}_{\mu\beta}[h,j] \right) = 0. \tag{4.113}$$

Two of the remaining equations inform us about the propagation of the two transverse-traceless modes of j obtainable from matching to the Interaction region. The relevant components for leading order propagation are,

$$\delta G^{(-2)}_{mm}[k] + \delta G^{(-1)}_{mm}[j] + 2\delta^2 G^{(-2)}_{mm}[h,j] = \frac{M}{\tilde{r}} \delta j''_{mm} - \frac{1}{2\tilde{r}} \delta j'_{mm} + \frac{1}{2} \partial_{\tilde{v}} j'_{mm} = 0, \tag{4.114a}$$

$$\delta G^{(-2)}_{\bar{m}\bar{m}}[k] + \delta G^{(-1)}_{\bar{m}\bar{m}}[j] + 2\delta^2 G^{(-2)}_{\bar{m}\bar{m}}[h,j] = \frac{M}{\tilde{r}} \delta j''_{\bar{m}\bar{m}} - \frac{1}{2\tilde{r}} \delta j'_{\bar{m}\bar{m}} - \frac{1}{2} \partial_{\tilde{v}} j'_{\bar{m}\bar{m}} = 0, \tag{4.114b}$$

The equations (4.114) enforce the slight corrections to simple $1/\tilde{r}$ propagation in the Far Zone. To further illustrate this point, consider a $Y_{\ell m}$ mode $j^{\ell m}_{mm}(\Theta, \tilde{r}, \tilde{v}) = \sum_k e^{ik\Omega(\tilde{u})\Theta} R_k(\tilde{r}, \tilde{u})$, where $R_k(r_0, \tilde{u})$ at some inner radius is constrained by matching to the Interaction Zone result. The equation (4.114a) then implies,

$$\frac{-ik\Omega(\tilde{u})M}{\tilde{r}} R_k(\tilde{r}, \tilde{u}) + \frac{1}{2\tilde{r}} R_k(\tilde{r}, \tilde{u}) + \frac{1}{2} R'_k(\tilde{r}, \tilde{u}) = 0, \tag{4.115}$$

which can be directly solved, yielding,

$$R_k(\tilde{r}, \tilde{u}) = \frac{R(r_0, \tilde{u}) r_0}{\tilde{r}} \tilde{r}^{2ik\Omega(\tilde{u})M}. \tag{4.116}$$

4.5.9 | Refinement

The formalism described so far breaks down at sufficiently large distances, $\tilde{r} \sim M/\varepsilon$ ($r \sim M/\varepsilon^2$) where it is no longer formally valid. To avoid this breakdown we now describe a refinement of the formalism that extends the domain of validity all the way out to future null infinity.

We can formally replace η with $\eta + \varepsilon h$, and replace h by zero, everywhere in the formalism. This yields a set of equations that are valid to the same order in ε as before. This requires reinterpreting all the covariant derivatives, parallel transport, inner products etc. to be those associated with the new background metric $\eta + \varepsilon h$, the Minkowski metric together with the linearized Kerr monopole. One can also to the same order take the background metric to be the Schwarzschild metric. In rescaled Schwarzschild coordinates $(\tilde{t}, \tilde{r}, \theta, \phi)$, the rescaled metric (4.79) can be written as

$$\varepsilon^2 ds^2 = -f d\tilde{t}^2 + \frac{1}{f} d\tilde{r}^2 + r^2 d\Omega^2, \tag{4.117}$$

with $f = 1 - 2\tilde{\varepsilon}M/\tilde{r}$. Here one is formally supposed to expand the dependence of w on $\tilde{\varepsilon}$ according to the original formalism, but we choose not to in order to obtain higher accuracy. We introduce the symbol $\tilde{\varepsilon} = \varepsilon$ in order to distinguish between the dependence on ε which we will expand out, and the dependence on $\tilde{\varepsilon}$ which we will not.

The result of this refinement is that we obtain the standard geometric optics formalism on the Schwarzschild background: The geodesic equation (4.103) becomes the geodesic equation in Schwarzschild, with the solution

$$\vec{l} = \frac{1}{f} \partial_{\tilde{t}} + \partial_{\tilde{r}}. \tag{4.118}$$

The corresponding phase function is $\Theta = -\tilde{u} = \tilde{r} + 2\tilde{\varepsilon}M\ln[\tilde{r}/(2\tilde{\varepsilon}M) - 1] - \tilde{t} = \varepsilon(r_* - t)$. The remaining components of the orthonormal basis can be taken to be

$$\vec{n} = \frac{1}{2}\partial_{\tilde{t}} - \frac{f}{2}\partial_{\tilde{r}}, \tag{4.119a}$$

$$\vec{m} = \frac{1}{\sqrt{2}r}\partial_\theta + \frac{i}{\sqrt{2}\sin(\theta)r}\partial_\phi, \tag{4.119b}$$

$$\vec{\bar{m}} = \frac{1}{\sqrt{2}r}\partial_\theta - \frac{i}{\sqrt{2}\sin(\theta)r}\partial_\phi, \tag{4.119c}$$

which is a refinement of the Minkowski orthonormal basis (4.108). The evolution equation (4.113) is then modified by dropping the third term, and has a simple $\propto 1/\tilde{r}$ solution, but now with the additional information that the definitions of $\tilde{u}$ and $\tilde{r}$ are refined.

4.5.10 | Discussion

We will present a full treatment of the geometric optics equations and the solutions for asymptotic and matching metric perturbations in our future paper [345]. There we will show that the domain of validity extends down to the edge of the far zone, $r \gg M$, and out to future null infinity for the refined formalism, by computing the post geometric optics correction $\delta k_{\alpha\beta}$ and determining when its contribution becomes comparable to that of $\delta j_{\alpha\beta}$.

In matching the far zone to the interaction zone, at first order information only flows out of the interaction zone to the far zone, by the first order metric perturbation fixing outgoing homogeneous modes. At subleading order, the backreaction j_0 generated quasistatic modes sourced by the first order perturbations in the far zone must be matched back into the interaction zone.

The set of Einstein equations that actually must be solved depends somewhat on the information we wish to extract from the inspiral. The tensor $\delta j_{\alpha\beta}$ describes the high frequency portion of the waveform, with frequencies $\sim M^{-1}$, while the tensor $j_{0\,\alpha\beta}$ describes a low frequency component with frequencies of order $\sim \varepsilon M^{-1}$, also known as the gravitational wave memory contribution. If the high frequency part of the post-adiabatic waveform is sufficient, and the dissipative part of the second-order self force is derived purely from interaction zone quantities, then one can simply use the leading order solutions to (4.103) whose coefficients are fixed by matching to the interaction zone. However, if we are interested in the gravitational wave memory portion of the waveform, or

in computing dynamical orbital invariants such as redshifts, we must solve the Einstein equation (4.101) to compute the backreaction of the emitted radiation. This computation is given in 6.

Finally, if we wish to optimize a second order computation by relying on fluxes rather than explicit evaluation of the second order dissipative self force near the small object, we must solve for the subleading radiation $\delta k_{\alpha\beta}$ using the rapidly varying part of Eq. (4.101). The initial results for the wave equations for δk and j_0 necessary for some of the post-adiabatic strategies, as well as an alternative Post-Minkowski description are detailed in Chapter 6.

4.6 The Near-Horizon Zone

For large negative values of the tortoise coordinate $-r^* \ll M$, a near-horizon expansion of the Einstein field equations is required. Much like the Far Zone, the leading order field equations should simply be the homogeneous wave equations, and therefore be fixed by the matching from the leading order Interaction Zone $h^{(1)}$. The homogeneous solutions in the Near-Horizon Zone will act as a source for second order modes, which will need to be calculated and matched to the second order interaction zone metric. The procedure for determining the Near-Horizon modes follows closely [302], which introduces an iterative technique for obtaining approximate solutions in Schwarzschild. The full details of the matching at the horizon and the implications for the Interaction Zone metric and worldline evolution will be given in the future paper [290].

4.7 Adiabatic order waveforms

4.7.1 Summary of results

By inserting the ansatz (4.16) into Einstein's equations, one obtains a set of equations that determines the free functions, order by order. At leading order we obtain an equation of the form

$$\mathcal{D} g^{(0)}_{\alpha\beta} = 0, \tag{4.120}$$

where $\mathcal{D}$ is a linear differential operator on the six dimensional manifold with coordinates $(f_r, f_\theta, \psi_\phi, \bar{x}^j)$. In solving this equation, $\tilde{w}$ is treated as a constant. The solution that matches

appropriately onto the worldline source can be written as

$$g^{(1)}_{\alpha\beta} = \frac{\partial g^{(0)}_{\alpha\beta}}{\partial M}\delta M(\tilde{w}) + \frac{\partial g^{(0)}_{\alpha\beta}}{\partial a}\delta a(\tilde{w}) + \mathcal{F}_{\alpha\beta}[\tilde{w}, q_i, \bar{x}^j, P^i(\tilde{w})]. \tag{4.121}$$

The first two terms in (4.121) describe the secular components, and are determined by the post-adiabatic equations of motion. They correspond to the slow accretion of energy and angular momentum by the central black hole over the course of the inspiral.

The final term in (4.121) has both secular and oscillatory pieces. Due to the angular momentum and energy of the orbiting body itself, there is an adiabatic-order correction to the zero-frequency component of the metric. We may choose $\mathcal{F}_{\alpha\beta}$ to be purely oscillatory either entirely within the inspiral orbit $r < r_{\rm min}$, or entirely outside the inspiral orbit $r > r_{\rm max}$. We choose $\mathcal{F}_{\alpha\beta}$ to be purely oscillatory at $r < r_{\rm min}$, so that the near-horizon time-averaged metric perturbation is simply the first two terms of (4.121). Define the time-averaged piece of the worldline-sourced metric perturbation $\bar{\mathcal{F}}_{\alpha\beta} = \langle \mathcal{F}_{\alpha\beta} \rangle$ and the purely oscillatory $\tilde{\mathcal{F}}_{\alpha\beta} = \mathcal{F}_{\alpha\beta} - \bar{\mathcal{F}}_{\alpha\beta}$

4.7.2 | Teukolsky equation at adiabatic order

Adiabatic waveform snapshots from a geodesic source are conveniently computed from the Teukolsky equation, which is equivalent to the linearized Einstein equation up to the nonradiative degrees of freedom. The Teukolsky formalism can similarly be adapted to a multiscale expansion, and here we highlight a few steps involved. First, using the metric (4.16), we obtain a similar expansion of the Weyl tensor. Projecting all quantities onto a Newman Penrose tetrad and following the usual derivations leads at the leading order to a differential equation that is similar to the standard Teukolsky equation. However, as before, the differential operators involved are operators on the larger, 6-dimensional manifold at fixed $\tilde{w}$. As a consequence, coordinate-time derivatives in the Teukolsky equation are replaced by $\partial_w \to \Omega^{A(0)}\partial_{q^A}$, at the leading order.

The second line of Eq. (4.121) is the oscillatory piece of the solution. Here, as we discuss in more detail below, one obtains a solution by taking the function $\mathcal{F}_{\alpha\beta}$ to be the function $\mathcal{F}_{\alpha\beta}(q^A, x^i, E, L_z, Q)$ that one obtains from standard linear perturbation theory with a geodesic source. This function is known analytically in Boyer-Lindquist coordinates (t, r, θ, ϕ) in terms of a mode expansion. For the remainder of this section, we discuss this mode expansion in terms of the

frequently used Teukolsky formalism, in which the dynamical variables are the first order corrections to Weyl scalars Ψ_0 and Ψ_4, which vanish in the Kerr background.

The leading order Teukolsky master variables, related to the Weyl scalars Ψ_0 and Ψ_4 satisfy an equation of the form

$$ {}_s\mathcal{O}^{(0)}\ {}_s\Psi^{(1)}(q^A,\tilde{w},x^i) = {}_s\mathcal{T}^{(1)}(q^A,\tilde{w},x^i), \tag{4.122}$$

where s is the spin weight ($+2$ for Ψ_0 and -2 for $(\Psi_2^{(0)})^{-2/3}\Psi_4$), $\mathcal{O}$ is a differential operator and ${}_s\mathcal{T}^{(1)}$ is a source term obtained from the stress-energy tensor. The solution will be of the form

$$ {}_s\Psi^{(1)}_{k_A lm} = {}_sR(r){}_s\Theta(\theta)e^{im\phi}e^{-ik_Aq^A}. \tag{4.123}$$

Substituting the ansatz (4.123) into the homogeneous version of Eq. (4.122) and keeping only the leading order terms results in the two equations:

$$\begin{aligned} 0 = \frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\ {}_s\Theta}{d\theta}\right) &+ \left[a^2\left(k_A\Omega^{A(0)}\right)^2\cos^2\theta - \frac{m^2}{\sin^2\theta} - 2ak_A\Omega^{A(0)}s\cos\theta\right. \\ &\left. - \frac{2ms\cos\theta}{\sin^2\theta} - s^2\cot^2\theta + \lambda - |s|\right]{}_s\Theta, \end{aligned} \tag{4.124}$$

$$\begin{aligned} 0 = \frac{1}{\Delta^s}\frac{d}{dr}\left(\Delta^{s+1}\frac{d\ {}_sR}{dr}\right) + &\left[\frac{(K^{(0)}_{mk_A})^2 - 2is(r-M)K^{(0)}_{mk_A}}{\Delta} + 4isk_A\Omega^{A(0)}r - \lambda\right. \\ &\left. - a^2\left(k_A\Omega^{A(0)}\right)^2 + 2amk_A\Omega^{A(0)} + s + |s|\right]{}_sR. \end{aligned} \tag{4.125}$$

Here, $\lambda_{slm}(ak_j\Omega^{(0)}_j)$ is the separation constant and we have defined

$$ K^{(0)}_{mk_A} = k_A\Omega^{A(0)}\varpi^2 - am. \tag{4.126}$$

The solutions to Eq. (4.124) are the real functions ${}_s\Theta_{lm}(ak_A\Omega^{A(0)},\theta)$ that are regular on $[0,\pi]$. In what follows, we do not show the dependence of ${}_s\Theta_{k_Alm}(\theta)$ on $ak_A\Omega^{A(0)}$ explicitly.

4.7.3 | Basis of modes

The radial equation (4.125) can be decomposed into a basis of modes with different asymptotic behavior analogous to what is done in standard black hole perturbation theory as summarized in the scalar case in [266] except for the following properties:

1. The scattering, transmission and normalization coefficients depend on the slow variable $\tilde{w}$, i. e. they are constant only at fixed $\tilde{w}$.

2. They depend on the harmonics $k_A\Omega^{A(0)}$ rather than ω.

Similarly, the complete Teukolsky mode functions can be obtained by replacing in the usual expressions the dependence $e^{-i\omega w}$ by $e^{-ik_A q^A}$ and using the generalized mode functions.

The retarded, advanced, and radiative Green's functions are constructed as in the standard formalism with the small modifications discussed above. From these expression for the Green's function, we can compute the field ${}_s\Psi(q^A, x^i, \tilde{w})$ generated by the source ${}_s\mathcal{T}(q^A, x^i, \tilde{w})$. All of these expressions depend on the amplitudes ${}_sZ^{\rm out/down}_{k_A lm}(\tilde{w})$ or ${}_sZ^{\rm in/up}_{k_A lm}(\tilde{w})$ whose properties we discuss next.

4.7.4 | Amplitudes

In this subsection, we show that the amplitudes ${}_sZ_{k_A lm}$ contain a term $\delta_{k_\phi m}$ and thus there are only four independent indices (k_r, k_θ, l, m) just as in the standard formalism. From the treatment of the orbital motion, it follows that the orbital phase $\phi(w)$ can be written as

$$\phi(w) = q^\phi + \sum_{k_A} \Phi_{k_A} e^{ik_A q^A} + O(\varepsilon) \equiv q^\phi + \delta\phi(q^a, \tilde{w}), \tag{4.127}$$

where we use the notation $q^a = (q^r, q^\theta)$ and $k_a = (k_r, k_\theta)$. The particle's stress-energy tensor is given by

$$T^{(1)}_{ab} = \mu \frac{u_a u_b}{\sqrt{-g}} \left(\frac{dw}{d\tau}\right)^{-1} \delta\left(r - r(q^a, \tilde{w})\right) \delta\left(\theta - \theta(q^a, \tilde{w})\right) \delta\left(\phi - \phi(q^\phi, q^a, \tilde{w})\right). \tag{4.128}$$

Here,

$$u_a = \left[-E^{(0)}(\tilde{w}), u^{(0)}_r[q^a, \tilde{w}], u^{(0)}_\theta[q^a, \tilde{w}], L^{(0)}_z(\tilde{w})\right], \tag{4.129}$$

and

$$\frac{dw}{d\tau} = \Gamma + \sum ik_a \Omega^{a(0)} T_{k_a}(\mathcal{J}^{(0)}_\lambda) \exp[ik_a q^a] + O(\varepsilon). \tag{4.130}$$

The amplitudes ${}_sZ^{\rm out}_{k_A lm}$ are computed by integrating the homogeneous solutions for certain complete mode functions ${}_s\Phi(q^A, x^i, \tilde{w})$ of the Hertz potential against the source obtained from acting with a

certain operator on (4.128). This leads to a result of the form

$$ {}_sZ^{\rm out}_{k_A lm}(\tilde{w}) = \frac{\mu}{(2\pi)^3}\int d^2q^a \int dq^\phi \int d^3x\ {}_{-s}R^{{\rm up}\ *}_{k_A lm}[r(q^a,\tilde{w})]\ {}_s\Theta_{k_A lm}[\theta(q^a,\tilde{w})]e^{im\phi(q^A,\tilde{w})} $$
$$ \times e^{-ik_\phi q^\phi - ik_a q^a}\mathcal{S}(q^a,\tilde{w}) \tag{4.131} $$
$$ = \frac{\mu}{(2\pi)^2}\int d^2q^a \int d^3x\ {}_{-s}R^{{\rm up}\ *}_{k_A lm}\left[r(q^a,\tilde{w})\right]\ {}_s\Theta_{k_A lm}\left[\theta(\psi_a,\tilde{w})\right] $$
$$ \times e^{im\delta\phi(q^a,\tilde{w})}e^{-ik_a q^a}\mathcal{S}(q^a,\tilde{w})\ \delta_{k_\phi\, m} \tag{4.132} $$

where in the second line we have used (4.127) and performed the integral over q_ϕ. Also, we denote $\mathcal{S}(q^a,\tilde{t}) = {}_s\tau_{ab}u^a u^b$, where ${}_s\tau_{ab}$ is an operator that takes the source of the Einstein equation to that of the Teukolsky equation.

4.7.5 | Waveforms

For $r\to\infty$, the quantity $(\Psi_2^{(0)})^{-2/3}\Psi_4$ is related to $h^{(1)}_{ab}$ by

$$ \psi_4^{(1)} = \frac{1}{2}\left(\Omega^{A(0)}\partial_{q^A}\right)^2\left(h_+^{(1)} - ih_\times^{(1)}\right). \tag{4.133} $$

The waveform can then be computed from

$$ h_+^{(1)} - ih_\times^{(1)} = 2\mathcal{I}_{\Omega^{A(0)}}\mathcal{I}_{\Omega^{A(0)}}((\Psi_2^{(0)})^{-2/3}\Psi_4), \tag{4.134} $$

where, for any function $F(\psi_A,\tilde{t})$, we define the anti-derivative operator

$$ \mathcal{I}_{\Omega^{A(0)}} = \sum_{k_A\neq 0}\frac{F_{k_A}}{ik_A\cdot\Omega^{A(0)}}e^{ik_A q^A}, \tag{4.135} $$

where $F_{k_A} = \int d^3q^A e^{-ik_A q^A}F(q^A)/(2\pi)^3$ are the Fourier coefficients of F.

From the expansions above, we obtain the explicit formula for the radiative fields

$$ h^{(1){\rm rad}}(x^i,q^A,\tilde{w}) = \sum_{k_a,l,m}\frac{1}{\left(k_a\Omega^{a(0)}+m\Omega^{\phi(0)}\right)^2}\left[\frac{\gamma^{\rm out}_{k_a lm}(\tilde{w})}{\alpha^*_{2k_a lm}(\tilde{w})}\rho(q^a,\tilde{w})^4\,{}_{-2}\Psi^{\rm out}_{k_a lm}(q^a,\tilde{w},x^i) + \text{"down"}\right]. \tag{4.136} $$

Here, $\rho = 1/(r-ia\cos\theta)$, and the coefficients $(\gamma_{k_a lm},\alpha_{k_a lm})$ are given by replacing in the usual results all constant coefficients with functions of $\tilde{w}$. The retarded fields are given by a very similar expression involving the "up" and "in" modes.

In the limit $r \to \infty$ the asymptotic behavior of the radial mode functions that are outgoing at $\mathcal{I}^+$ is

$$\lim_{r\to\infty} {}_{-2}R^{\rm up} \to \beta_{-2k_a lm} \mid k_a \Omega^{a(0)}(\tilde{w}) + m\Omega^{\phi(0)}(\tilde{w}) \mid^{-1/2} \; r^3 \; e^{i\left[k_a \Omega^{a(0)} + m\Omega^{(0)}_{\phi}\right] r^*}. \tag{4.137}$$

Using also that in this limit $\rho^4 \to r^{-4}$, the leading order retarded waveform at $r \to \infty$ has the asymptotic behavior

$$\left(h^{(1)}_{+} - ih^{(1)}_{\times}\right)^{\infty} \sim \frac{1}{r} \sum_{k_a=-\infty}^{\infty} \sum_{l=2}^{\infty} \sum_{m=l}^{l} \frac{1}{\left(k_a \Omega^{a(0)}(\tilde{w}) + m\Omega^{\phi(0)}(\tilde{w})\right)^2} \frac{\gamma_{k_A lm}(\tilde{w})}{|k_a \Omega^{a(0)}(\tilde{w}) + m\Omega^{\phi(0)}(\tilde{w})|^{1/2}} \tag{4.138}$$

$$\times \, {}_{-2}S_{k_A lm}(\theta,\phi) e^{-ik_A\left[q^A(\tilde{w}) - \Omega^{A(0)} r^\star\right]} \tag{4.139}$$

with $k_\phi = m$. This shows that at this order, no matching at large r is required to read off the asymptotic waveform.

4.8 | Conclusions

We have introduced the extension of the multiscale framework [222] to the Einstein field equations. The adiabatic order computation presented in this paper is sufficient to obtain the inspiral worldline and amplitude and frequency of the asymptotic waveform with only $\mathcal{O}(\varepsilon)$ error. To this precision, the results closely resemble the method of osculating geodesics. In this procedure, we anticipate that the non-conserved geodesic parameters are evolved in slow time $\tilde{w}$ using balance laws, and the final waveform is derived from evaluating the relation between the multiscale time variables and the coordinate time. The result is intuitively accessible: the waveform instantaneously resembles the oscillatory waveform from a pure geodesic, but the amplitude and frequency slowly drift a the inspiral evolves. Notably, this excludes the majority of modern techniques in self force procedures, as much of the current technology of self force computations revolves around either first order conservative self force or second order dissipative self force, which both enter at post-adiabatic order. Importantly, *all* such effects must be included for a post-adiabatic waveform.

This article lays the foundation for our future publications, which will develop the formalism to post-adiabatic order. The results from the post-adiabatic computation will include an algorithm

for computing the phase of the orbit with at most $\mathcal{O}(\varepsilon)$ error, as well as improving the precision of the other important inspiral outputs. The full Post-adiabatic information of the inspiral can also provide a valuable check on the complimentary techniques of numerical relativity, for low mass ratios, and Post-Newtonian expansions, for large body separations.

Appendix

4.A Coordinate-time force definitions

Very similar to the derivation in [222], we'd like to find an explicit form of the acceleration term that should appear on the right-hand side of the Hamilton-Jacobi equations. To find this, we'll first write the single geodesic equation we wish the system to obey:

$$\frac{d}{dw}v^i + \frac{d^2\tau}{dw^2}v^i + \Gamma^i_{kl}v^k v^l + 2\Gamma^i_{kw}v^k + \Gamma^i_{ww} = a^i\left(-g_{ww} - 2g_{iw}v^i - g_{ij}v^i v^j\right), \tag{4.140}$$

where the right-hand side is scaled by the appropriate $(d\tau/dw)^2$ to permit our use of the coordinate time derivatives.

Recall that the Hamiltonian that describes the dynamics of the geodesic system using the coordinate time is written (4.61):

$$H = \frac{g^{wi}}{g^{ww}}p_i + \frac{1}{|g^{ww}|}\left((g^{wi}p_i)^2 - g^{ww}(g^{ij}p_i p_j + \mu^2)\right)^{1/2} \tag{4.141}$$

The Hamilton-Jacobi equations for the strict geodesic are then

$$\frac{dx^i}{dw} = v^i = \frac{g^{wi}}{g^{ww}} + \frac{1}{|g^{ww}|}\frac{g^{wi}g^{wj}p_j - g^{ww}g^{ij}p_j}{((g^{wi}p_i)^2 - g^{ww}(g^{ij}p_i p_j + \mu^2))^{1/2}} \tag{4.142a}$$

$$\begin{aligned}\frac{dp_i}{dw} = &-\frac{\partial_i g^{wj}p_j}{g^{ww}} + \frac{1}{(g^{ww})^2}\partial_i g^{ww}g^{wj}p_j + \frac{1}{(g^{ww})^2}\partial_i g^{ww}\left((g^{wl}p_l)^2 - g^{ww}(g^{lm}p_l p_m + \mu^2)\right)^{1/2} \\ &+\frac{1}{g^{ww}}\frac{\partial_i g^{wj}p_j g^{wk}p_k - \partial_i(g^{ww}g^{jk})p_j p_k}{((g^{wi}p_i)^2 - g^{ww}(g^{ij}p_i p_j + \mu^2))^{1/2}}.\end{aligned} \tag{4.142b}$$

The first of these equations defines the conversion between the coordinate time three-velocity and the conjugate momentum. The second of these equations will reproduce the geodesic equations, without the additional acceleration term. We will impose an additional f_i term on the right-hand side of the momentum equation to ensure that it agrees with the forced geodesic equation (4.140).

Using our value for v^i, we can derive the expression for dv^i/dw in terms of p_j:

$$\begin{aligned}\frac{d}{dw}v^i = \frac{1}{|g^{ww}|}&\left((g^{wl}p_l)^2 - g^{ww}(g^{lk}p_lp_k+\mu^2)\right)^{-3/2}\\ &\times\Big(((g^{wl}p_l)^2 - g^{ww}(g^{lk}p_lp_k+\mu^2))(g^{wi}g^{wj}\dot{p}_j - g^{ww}g^{ij}\dot{p}_j)\\ &\quad-(g^{wi}g^{wl}p_l - g^{ww}g^{il}p_l)(g^{wk}p_kg^{wm}\dot{p}_m - g^{ww}g^{km}p_k\dot{p}_m)\Big)\end{aligned} \tag{4.143}$$

Setting this equal to the remaining terms in (4.140), and inverting for the additional forcing term required in the momentum equation, we obtain:

$$f_i = \mu\frac{d\tau}{dw}\,(a^\alpha g_{\alpha i}) = \mu\frac{d\tau}{dw}a_i, \tag{4.144}$$

where

$$\left(\frac{dw}{d\tau}\right)^2 = \frac{1}{\mu^2}((g^{wk}p_k)^2 - g^{ww}(g^{jm}p_jp_m+\mu^2)). \tag{4.145}$$

The forced geodesic equations may now be written:

$$\frac{dx^i}{dw} = v^i = \frac{g^{wi}}{g^{ww}} - \frac{1}{\mu\, g^{ww}}\frac{d\tau}{dw}\left(g^{wi}g^{wj}p_j - g^{ww}g^{ij}p_j\right) \tag{4.146a}$$

$$\begin{aligned}\frac{dp_i}{dw} = &-\partial_i\left(\frac{g^{wj}}{g^{ww}}\right)p_j + \frac{\mu}{(g^{ww})^2}\frac{dw}{d\tau}\partial_i g^{tt}\\ &+\frac{1}{\mu g^{ww}}\frac{dw}{d\tau}\left(\partial_i g^{wj}p_jg^{wk}p_k - \partial_i(g^{ww}g^{jk})p_jp_k\right) + f_i(h,g,p).\end{aligned} \tag{4.146b}$$

From the forced geodesic equations in terms of the original variables (x,p), we can use the definitions of the action-angle variables in coordinate time (4.62,4.67) to derive the form of the forcing terms that we include in the right-hand side of our multiscale equations of motion

$$\frac{dq_i}{dw} = \omega_i(J_j) + \varepsilon g_i^{(1)}(q_r,q_\theta,J_j) + \varepsilon^2 g_i^{(2)}(q_r,q_\theta,J_j) + \mathcal{O}(\varepsilon^3) \tag{4.147}$$

$$\frac{dJ_i}{dw} = \varepsilon G_i'^{(1)}(q_r,q_\theta,J_j) + \varepsilon^2 G_i'^{(2)}(q_r,q_\theta,J_j) + \mathcal{O}(\varepsilon^3). \tag{4.148}$$

From the use of the canonical transformation $(x,p)\to(q,J)$, we have:

$$G_k' = \frac{\partial J_k}{\partial P_j}\left(\frac{\partial P_j}{\partial p_i}\right)_x f_i \tag{4.149a}$$

$$g_k = \left(\frac{\partial q_k}{\partial p_i}\right)_x f_i, \tag{4.149b}$$

Due to the implicit of the angle variables on the momentum values via the generator of the canonical transformation, $\mathcal{W}$ (4.65), the explicit expression for the angle forcing terms g_k requires additional expansion [222]:

$$g_k = a_i \left(\frac{\partial P_l}{\partial p_i}\right)_x \frac{\partial P_j}{\partial J_i}\left[\left(\frac{\partial^2 \mathcal{W}}{\partial P_j \partial P_l}\right)_x - \left(\frac{\partial \mathcal{W}}{\partial P_m}\right)_x \frac{\partial P_m}{\partial J_k}\frac{\partial^2 J_k}{\partial P_j \partial P_l}\right] \tag{4.150}$$

We can use the results for the first derivative of the momentum to verify the form of the first derivatives of $\{E, L_z, K\}$ given in section III. The time derivative of the energy of the orbit, which is equivalent to the new Hamiltonian (4.61) is

$$\frac{dE}{dw} = \mu\left(\frac{g^{wi}}{g^{ww}}f_i - \frac{1}{g^{ww}}\frac{d\tau}{dw}\frac{1}{\mu}\left(g^{wi}p_i g^{wj}f_j - g^{ww}g^{kl}p_l f_l\right)\right). \tag{4.151}$$

Now, we use the identity $a^\mu p_\mu = 0$ to write:

$$\begin{aligned} a_w = -\frac{1}{\mu}\frac{d\tau}{dw}a_i p^i &= -\frac{1}{\mu}\frac{d\tau}{dw}a_k\left(g^{kl}p_l + g^{kw}p_w\right) \\ &= -\frac{1}{\mu}\frac{d\tau}{dw}a_k\left(g^{kj}p_j + \frac{g^{kw}}{g^{ww}}p^w - \frac{g^{kw}g^{wj}}{g^{ww}}p_j\right), \end{aligned} \tag{4.152}$$

which implies the simplification of the energy equation,

$$\frac{dE}{dw} = -\frac{d\tau}{dw}a_w. \tag{4.153}$$

The corresponding equation for $L_z = p_\phi$ is immediate:

$$\frac{dL_z}{dw} = \mu\frac{d\tau}{dw}a_\phi \tag{4.154}$$

4.B | Hyperboloidal time coordinate

Let us briefly expand on how the expansion would proceed, and ultimately fail, under spacelike slices of constant time coordinate t. A slow time variable assigned as $\varepsilon t_{\rm sch} = \tilde{t}_{\rm sch}$ by scaling the ordinary Schwarzschild time causes problems at large scales. Expanding the multiscale equations via the ansatz using $\tilde{t}_{\rm sch}$ has the direct mathematical consequence of a divergent second-order field. This sort of divergence is discussed in detail in [278]; here we will present a rough scaling argument to justify our choice of time coordinate.

Consider the wave equation at long distances from the source. Under the derivative expansion (4.12), we will obtain a second order source from the $\mathcal{O}(\varepsilon)$ time derivatives in the first-order wave equation. We will consider both the case in which the slow time variable is asymptotically equal to $\varepsilon t_{\rm sch}$ and the case where the slow time variable $\tilde{w}$ is asymptotically equal to the scaled retarded time εu at large radius $r^* \gg M/\varepsilon$.

The choice of slow time variable gives rise to distinct source terms at second order. With a $\tilde{t}_{sch}$ slow time, the radial homogeneous modes will behave as $\propto \frac{e^{ik_A \Omega^A(\tilde{t}) r}}{r}$, while the same wave expressed with a retarded $\tilde{u}$ slow time scales as simply $\frac{1}{r}$ asymptotically. The additional source term obtained from (4.12) in terms of Schwarzschild slow time $\tilde{t}_{\rm sch}$ behaves as $\propto r^0$ asymptotically, which would cause unacceptable divergence in the second order solution. Conversely, when the source term for the equation is expressed in terms of the slow time $\tilde{u}$, it scales as $1/r^2$, which grants a solvable wave equation.

We will avoid imposing any particular choice of coordinate time, and instead impose Boyer-Lindquist coordinates $\{r, t', \theta, \phi\}$ on the background metric, then shift the time variable by:

$$w = t' + h(r) \qquad \left| \begin{array}{l} \lim_{r\to\infty} h(r) = -r \\ \lim_{r\to 2M} h(r) = r \end{array} \right. \tag{4.155}$$

The metric is then adjusted from Boyer Lindquist as:

$$\begin{aligned} ds^2 = &- \left(1 - \frac{2Mr}{\Sigma}\right) dw^2 + 2\left(1 - \frac{2Mr}{\Sigma}\right) h'(r) dr dw \\ &+ \left(-\left(1 - \frac{2Mr}{\Sigma}\right) h'^2(r) + \frac{\Sigma}{\Delta}\right) dr^2 + \Sigma d\theta^2 + \left(r^2 + a^2 + \frac{2Ma^2 r}{\Sigma}\sin^2\theta\right)\sin^2\theta d\phi^2 \\ &+ \frac{4Mar}{\Sigma} h'(r) \sin^2\theta dr d\phi - \frac{4Mar}{\Sigma}\sin^2\theta dw d\phi \end{aligned} \tag{4.156}$$

4.C | Near-particle matching details

The multiscale formalism requires a support computation to determine the information about how the small companion will respond to the multiscale metric perturbation (the self force), and about the backreaction of the small companion on the metric perturbation. We have developed two methods for obtaining this information

4.C.1 | Expansion near fixed $\tilde{w}$

In this approach, we accept the limitation of the self-consistent approximation that it cannot hold for times of the order of the inspiral $w \sim M/\varepsilon$. Instead, for the step at which we evolve the slowly varying multiscale quantities, we rely on the self consistent results, evaluated for the Multiscale metric perturbation expanded as $\tilde{w} = \tilde{w}_0 + \varepsilon w$. In this expansion, we assume that $w \sim M$, so a re-expansion in the explicit ε is justified. This computation will then need to be repeated for each $\tilde{w}_0$ over the course of the inspiral.

In principle, the self force and near-companion puncture we suggest here could be derived from [249]. However, the results in that publication have subtle dependence on high powers of the coordinate time, associated with the gauge choices made for the computation. It should be possible to perform the same self force and puncture computation in a gauge more compatible with the multiscale formalism, and in that case we expect the resulting equations would be consistent with what we present here. This suspicion is motivated by the computation performed in [303] of the self force in a Gralla-Wald form.

Here we only sketch the form of the computation, and appeal to the comparison to the steps in the hybrid section (below) to justify the final form of the self-acceleration and puncture. We assume the multiscale metric expansion

$$g_{\mu\nu} = g^{(0)}_{\mu\nu}[x^i] + \varepsilon h^{(1)}_{\mu\nu}[\varphi^A, \tilde{w}, x^i] + \varepsilon^2 h^{(2)}_{\mu\nu}[\varphi^A, \tilde{w}, x^i] + \mathcal{O}(\varepsilon^3) \tag{4.157}$$

To evaluate about a fixed slow time, the $\tilde{w}$ dependence may be evaluated directly. The implicit time dependence of the fast-time variable must also be expanded about the fixed time

$$\varphi^A(\tilde{w}_0 + \varepsilon t) = \varphi^A(\tilde{w}_0) + w\Omega^{A(0)}(\tilde{w}_0) + \varepsilon w\Omega^{A(1)}(\tilde{w}_0) + \varepsilon w^2\Omega'^{A(0)}(\tilde{w}_0) + \mathcal{O}(\varepsilon^2) \tag{4.158}$$

The metric perturbation is then given as a function of $\{\tilde{w}_0, t, x^i\}$:

$$g_{\mu\nu} = g^{(0)}_{\mu\nu}[x^i] + \varepsilon H^{(1)}[w, x^i] + \varepsilon^2 H^{(2)}[w, x^i] + \mathcal{O}(\varepsilon^3), \tag{4.159}$$

where the new metric perturbations $H^{(1)}$ and $H^{(2)}$ are defined as

$$\begin{aligned}
H^{(1)}[w, x^i] =& h^{(1)}\left[\varphi(\tilde{w}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0, x^i\right] \\
H^{(2)}[w, x^i] =& (w\Omega^{(1)A}(\tilde{w}_0) + w^2\Omega'^{(0)A})\partial_{\varphi^A} h^{(1)}\left[\varphi(\tilde{w}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0, x^i\right] \\
&+ w\partial_{\tilde{w}} h^{(1)}\left[\varphi(\tilde{w}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0, x^i\right] + h^{(2)}\left[\varphi(\tilde{w}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0, x^i\right]
\end{aligned} \tag{4.160}$$

The worldline function is expanded similarly, using its dependence on $\{\varphi^A, \tilde{w}\}$:

$$\begin{aligned}
z^\mu(t) =& z^{(0)\mu}\left[\varphi(\tilde{w}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0\right] \\
&+ \varepsilon\Big((w\Omega^{(1)A} + w^2\Omega'^{(0)A})\partial_{\varphi^A} z^{(0)\mu}\left[\varphi^A(\tilde{t}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0\right] \\
&\quad + w\partial_{\tilde{w}} z^{(0)\mu}[\varphi(\tilde{w}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0] + z^{(1)}[\varphi(\tilde{w}_0) + w\Omega^{(0)}(\tilde{w}_0), \tilde{w}_0]\Big)
\end{aligned} \tag{4.161}$$

The self-consistent formalism may similarly be expanded about a fixed time. The resulting acceleration is given in [303], evaluated at the metric perturbation similarly expanded about a fixed time. The first order acceleration is evaluated first, subsequently allowing the slow time derivatives in the second-order expressions to be evaluated.

Finally, the puncture metric may also be expanded using the worldline's time dependence, obtaining the correct puncture metric as a function of $\{\tilde{w}_0, w, x^i\}$. Therefore, at each fixed $\tilde{w}_0$, we can extract the correct metric to match to the multiscale metric evaluated at the same time.

4.C.2 | Hybrid reasoning

Instead of the original Self-Consistent expansion, consider introducing, at each order, a $h'^{(n)}(\varepsilon w, x^i)$, which contains the secularly growing mass and spin of the central black hole. These additional contributions to the metric perturbation may be derived from the re-expansion of the multiscale ansatz, keeping the purely slow time $\tilde{w}$ dependent parts unchanged.

Once again, start with the multiscale metric ansatz

$$g_{\mu\nu} = g^{(0)}_{\mu\nu} + \varepsilon h^{(1)}_{\mu\nu}[\varphi^A, \tilde{w}, x^i] + \varepsilon^2 h^{(2)}_{\mu\nu}[\varphi^A, \tilde{w}, x^i] + \mathcal{O}(\varepsilon^3) \tag{4.162}$$

For the subclass of orbits z^μ which possess a multiscale expansion,

$$z^\mu[x^\nu] = z^{(0)\mu}[\varepsilon w, \varphi^A(\varepsilon w)] + \varepsilon z^{(1)\mu}[\varepsilon w, \varphi^A(\varepsilon w)] + \varepsilon^2 z^{(2)\mu}[\varepsilon w, \varphi^A(\varepsilon w)] + \mathcal{O}(\varepsilon^3), \tag{4.163}$$

we may define a self-consistent expansion using osculating geodesics.

Define new functions $P^M(z^\alpha, w)$:

$$P^M(z^\alpha, w) = \begin{cases} -K^t_\mu u^\mu(w) \\ K^\phi_\mu u^\mu(w) \\ K_{\mu\nu} u^\mu(w) u^\nu(w) \end{cases} \tag{4.164}$$

And functions $\varphi^A(z^\alpha, t, P^M)$:

$$\varphi^A = \varphi^A(z^\alpha, t, P^M) + \mathcal{O}(\varepsilon^2) = \frac{\partial \mathcal{W}(z^\alpha, t, J^A(P^M))}{\partial J^A(P^M)} + \mathcal{O}(\varepsilon^2), \tag{4.165}$$

where the notation for the action-angle variables is used from [222]. To be explicit, in the case of a Schwarzschild background, these functions could be written explicitly as,

$$\begin{aligned}
\varphi^t &= t + \int^r dr h'(r) \mp \int^r \frac{r^2 E}{r-2M}\left(-\left(1-\frac{2M}{r}\right)L_z^2 + r^2E^2 - r^2\left(1-\frac{2M}{r}\right)\mu^2\right)^{-1/2} + \mathcal{O}(\varepsilon^2) \\
\varphi^\phi &= \phi \mp \int^r \frac{1}{r} L_z \left(-\left(1-\frac{2M}{r}\right)L_z^2 + r^2E^2 - r^2\left(1-\frac{2M}{r}\right)\mu^2\right)^{-1/2} + \mathcal{O}(\varepsilon^2) \\
\varphi^r &= 4\pi \frac{\int^r r\left(-\left(1-\frac{2M}{r}\right)L_z^2 + r^2E^2 - r^2\left(1-\frac{2M}{r}\right)\mu^2\right)^{-1/2} dr}{\oint r\left(-\left(1-\frac{2M}{r}\right)L_z^2 + r^2E^2 - r^2\left(1-\frac{2M}{r}\right)\mu^2\right)^{-1/2}} dr,
\end{aligned} \tag{4.166}$$

where the w component, if needed, can easily be inferred from the φ^t and φ^r components.

Finally, we may use these functions to evaluate the original metric perturbation 'off-shell':

$$\begin{aligned}
&h^{(1)}_{\mu\nu}(\delta M(\tilde{w}), \delta a(\tilde{w}), x^i, \tilde{w}, \varphi^A(z^\alpha, w, P^M(w, z^\mu))) \\
&\qquad \equiv g^{\delta M}_{\mu\nu} \delta M^{(1)}(\tilde{w}) + g^{\delta a}_{\mu\nu} \delta a^{(1)}(\tilde{w}) + \mathcal{F}^{(n)}_{\mu\nu}\left(x^i, \varepsilon w, \varphi^A(\tilde{w})\right)\Big|_{\varphi^A = \varphi^A(z^\alpha, t, P^M(t, z^\mu))} \\
&\qquad \equiv h^{(1)SC}_{\mu\nu}[x^i, w, z^\mu] + h'^{(n)SC}_{\mu\nu}[x^i, \tilde{w}]
\end{aligned} \tag{4.167}$$

The later parts of the argument will also require the inverse transformation, for which we evaluate the functional argument of the metric perturbation using the multiscale worldline.

The multiscale formalism for the orbits is constructed using the one-to-one relation between orbital position for a set of geodesic parameters, and the action-angle variables holds for all fixed $\tilde{w}$, for the full multiscale worldline. Re-evaluating the h^{SC} on $z^\mu = z^{(0)\mu}[\varphi^A, \tilde{w}] + \varepsilon z^{(1)\mu}[\varphi^A, \tilde{w}]$, evaluates the frequency sum (4.10) to the maximum order of the perturbation.

$$\varphi^A(z^{\alpha(0)}(\tilde{w},\varphi^A)+\varepsilon z^{\alpha(1)}(\tilde{w},\varphi^A)+\mathcal{O}(\varepsilon^2))=\varphi^A+\mathcal{O}(\varepsilon), \tag{4.168}$$

so the inverse transformation proceeds transparently.

The metric expansion now takes the form

$$g_{\mu\nu}=g^{(0)}_{\mu\nu}+\varepsilon h^{(1)SC}_{\mu\nu}[x^i,w;z^\nu]+\varepsilon h'^{(1)SC}_{\mu\nu}(\tilde{w},x^i)+\varepsilon^2 h^{(2)SC}_{\mu\nu}[x^i,w;z^\nu]+\varepsilon^2 h'^{(2)SC}_{\mu\nu}(\tilde{w},x^i) \tag{4.169}$$

The motivation for the incomplete re-expansion (hybrid ansatz) above is to obtain a formalism which simultaneously benefits from the expression of the worldline on-shell constraint from the self-consistent formalism and the mass evolution constraint from the multiscale formalism. From this point forward, we will drop the SC from the superscript of these functions. They will be distinguished from the multiscale functions by their respective arguments.

Recall that the gauge condition may be imposed (for the Self Consistent formalism):

$$\nabla_\mu \bar{h}^{\mu\nu}_{\text{full}}=\nabla_\mu\left(\bar{h}^{\mu\nu}[x^i,w;z^\mu]+\bar{h}'^{\mu\nu}[\tilde{w},x^i]\right)=0, \tag{4.170}$$

The gauge condition can instead be imposed order-by-order, provided we appropriately expand the worldline ε dependence, which is the tactic used for the Multiscale formalism:

$$\nabla_\mu \bar{h}^{(1)\mu\nu}[x^i,\tilde{w},\varphi^A]+\nabla_\mu \bar{h}'^{(1)\mu\nu}[\tilde{w},x^i]=0 \tag{4.171a}$$

$$w_\mu\partial_{\tilde{w}}\bar{h}^{(1)\mu\nu}[x^i,\tilde{w},\varphi^A]+\partial_{\tilde{w}}\bar{h}'^{(1)t\nu}[\tilde{w},x^i]+\nabla_\mu \bar{h}^{(2)\mu\nu}[x^i,\tilde{w},\varphi^A]+\nabla_\mu \bar{h}'^{(2)\mu\nu}[\tilde{w},x^i]=0, \tag{4.171b}$$

These conditions can then be used to find the mass and spin evolution of the central black hole, as well as the perturbative expansion of the worldline, although the worldline expansion would require a full rederivation of the self-force, similar to [287], under this adjusted set of assumptions.

With the multiscale formalism, we expand a worldline function in powers of ε at fixed $\tilde{w}$, which permits a perturbative worldline expansion which preserves long-scale fidelity to the exact worldline.

In the hybrid scheme, we will assume that the computation has access to both types of worldline: a perturbatively expanded (multiscale) worldline $z^{(n)\mu}$, and a semi-exact worldline z^μ. The perturbatively expanded worldline will satisfy the multiscale assumption in the t variable; that is, it will be multiperiodic in the appropriate action-angle variables φ^A (also defined in terms of the worldline), and possess a slow time $\tilde{w}$ dependence. We may define the semi-exact worldline as the

sum to the desired order of approximation of the multiscale worldline:

$$z^\mu(w) = z^{(0)\mu}[\tilde{w}, \varphi^A(w, z^\mu)] + \varepsilon z^{(1)\mu}[\tilde{w}, \varphi^A(w, z^\mu)] + \mathcal{O}(\varepsilon^2) \tag{4.172}$$

Using the exact gauge condition, we are able (via [287]) to derive a perturbative expansion of the acceleration of the worldline, in terms of the regular part of the metric perturbation (defined by subtraction of the puncture).

$$\begin{aligned} a^\mu(w) = &\varepsilon a^{(1)\mu}(w, z^\mu, h^{\mathcal{R}(1)}[x^i, w, z^\mu]) \\ &+ \varepsilon^2 a^{(2)\mu}(w, z^\mu, h^{\mathcal{R}(1)}[x^i, w, z^\mu], h^{\mathcal{R}(2)}[x^i, w, z^\mu]) + \mathcal{O}(\varepsilon^3). \end{aligned} \tag{4.173}$$

The semi-exact acceleration would be used in a self-consistent approximation to evolve the semi-exact worldline, which could then be used with the field equations to generate the solution for the inspiral. By the assumption that the worldline has a multiscale expansion, the acceleration should also have a multiscale expansion, given by the explicit expression of the worldline in terms of the multiscale $\{\tilde{w}, \varphi^A\}$. As this proceeds transparently (described above) for the metric perturbation, the expression becomes:

$$\begin{aligned} a^\mu(\tilde{w}, \varphi^A) = &\varepsilon a^{(1)\mu}(t, z^{(0)\mu}[\tilde{w}, \varphi^A], h^{\mathcal{R}(1)}[x^i, \tilde{w}, \varphi^A]) \\ &+ \varepsilon^2 a^{(2)\mu}(t, z^{(0)\mu}[\tilde{w}, \varphi^A], h^{\mathcal{R}(1)}[x^i, w, \varphi^A], h^{\mathcal{R}(2)}[x^i, w, \varphi^A]) \\ &+ \varepsilon^2 \mathbb{d} a^{(1)\mu}(t, z^{(0)\mu}[\tilde{w}, \varphi^A], z^{(1)\mu}[\tilde{w}, \varphi^A], h^{\mathcal{R}}) + \mathcal{O}(\varepsilon^3). \end{aligned} \tag{4.174}$$

If this were to be developed numerically, the evaluation of the multiscale acceleration function could be obtained by a wavelet decomposition or similar local Fourier decomposition.

The result of this evaluation is the acceleration, derived in [287], expanded about a perturbative worldline, as in [381]:

$$\begin{aligned} a^{(1)\alpha} =& -\frac{1}{2}P^{\alpha\mu}(2h^{\mathcal{R}(1)}_{\delta\beta;\gamma} - h^{\mathcal{R}(1)}_{\beta\gamma;\delta})u^\beta u^\gamma \\ a^{(2)\alpha} =& -\frac{1}{2}P^{\mu\rho}\left(2h^{\mathcal{R}(2)}_{\rho\sigma;\lambda} - h^{\mathcal{R}(2)}_{\sigma\lambda;\rho}\right)u^\sigma u^\lambda - \frac{1}{2}P^{\mu\rho}\left(2h^{\mathcal{R}(1)}_{\rho\sigma;\lambda\delta} - h^{\mathcal{R}(1)}_{\sigma\lambda;\rho\delta}\right)u^\sigma u^\lambda z^{(1)\delta}_\perp \\ &- \left(2h^{\mathcal{R}(1)}_{\nu\sigma;\lambda} - h^{\mathcal{R}(1)}_{\sigma\lambda;\nu}\right)\left(\frac{1}{2}u^{(1)\mu}_\perp u^\nu u^\sigma u^\lambda + P^{\mu\nu}u^{(1)(\sigma}_\perp u^{\lambda)}\right) + \frac{1}{2}P^{\mu\nu}h^{\mathcal{R}(1)\rho}_\nu\left(2h^{\mathcal{R}(1)}_{\rho\sigma;\lambda} - h^{\mathcal{R}(1)}_{\sigma\lambda;\rho}\right)u^\sigma u^\lambda \\ &- \frac{1}{2}P^{\alpha\mu}\left(2t_\gamma\partial_{\tilde{w}}h^{\mathcal{R}(1)}_{\delta\beta} - t_\delta\partial_{\tilde{w}}h^{\mathcal{R}(1)}_{\beta\gamma}\right)u^\beta u^\gamma, \end{aligned} \tag{4.175}$$

where the superscript on the velocity refers to the expansion of the worldline $u^{(n)} \equiv dz^{(n)}/d\tau$, and the $\perp$ subscript indicates the component projected perpendicular to the velocity u^μ.

However, using the exact gauge condition alone, as described in the previous section, would neglect the slow evolution of the background spacetime. To correctly compute the slowly evolving spacetime mass and spin, we will need the right-hand side of (4.172).

In principle, the worldline function determines values for the frequencies $\Omega^A(\tilde{w})$. We may then uniquely re-express the acceleration functions in terms of multiscale parameters via a Fourier decomposition. We may therefore compute a set of multiscale accelerations $a^{(n)\mu}(\tilde{w}, \varphi^A)$.

Using these accelerations, we can simultaneously evolve the set of worldline functions on the right-hand side of (4.172). Finally, the slowly varying portion of the metric perturbation $h'[\tilde{w}, x^i]$ is determined by the quasistatic portion of the perturbative gauge condition, evaluated on the expanded, multiscale worldline at fixed $\tilde{w}$. The first nontrivial slow-time gauge condition is

$$w_\mu \partial_{\tilde{w}} \bar{h}^{(1)\mu\nu}[x^i, \tilde{w}, \varphi^A] + \nabla_\mu \bar{h}^{(2)\mu\nu}[x^i, \tilde{w}, \varphi^A] + \nabla_\mu \bar{h}'^{(2)\mu\nu}[\tilde{w}, x^i] + \partial_{\tilde{w}} \bar{h}'^{(1)t\nu}[\tilde{w}, x^i] = 0, \tag{4.176}$$

The first two of the terms on the left-hand side of this equation will be fixed at a particular $\tilde{w}$ by the wave and worldline evaluation steps. The third term will be constructed to precisely vanish - the slow-time evolving contributions will be constructed to satisfy the $\tilde{w}$-independent gauge condition. Finally, the last term is the undetermined quantity which will fix the slowly evolving parameters of the spacetime - in particular, the mass and spin of the central black hole.

The puncture metric, given in [250] is given in terms of Synge's worldfunction about the worldline. The puncture metric will also need to be expanded about the multiscale $z^{(n)\mu}$, which will include a dipole moment correction at subleading order. The value we require from this puncture metric expansion is the first subleading dipole correction to the first-order puncture metric, which is a second-order quantity. We will not need to adjust the second order puncture expressions from [250] at all, as any such corrections are at least third order in ε.

Chapter 4 Bibliography

[198] B. P. Abbott *et al.*, *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116** (2016) (6), p. 061102

[199] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, E. S. Phinney and M. Vallisneri, *Event rate estimates for LISA extreme mass ratio capture sources*, Class. Quant. Grav. **21** (2004), pp. S1595–S1606

[200] L. Blanchet, *Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries*, Living Rev. Rel. **17** (2014), p. 2

[201] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna and D. Shoemaker, *Georgia Tech Catalog of Gravitational Waveforms*, Class. Quant. Grav. **33** (2016) (20), p. 204001

[202] A. H. Mroue *et al.*, *Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy*, Phys. Rev. Lett. **111** (2013) (24), p. 241104

[203] M. Shibata and K. Taniguchi, *Coalescence of Black Hole-Neutron Star Binaries*, Living Reviews in Relativity **14** (2011) (6),
URL: http://www.livingreviews.org/lrr-2011-6

[204] J. A. Faber and F. A. Rasio, *Binary Neutron Star Mergers*, Living Reviews in Relativity **15** (2012) (8),
URL: http://www.livingreviews.org/lrr-2012-8

[205] A. Buonanno and T. Damour, *Effective one-body approach to general relativistic two-body dynamics*, Phys. Rev. D **59** (1999), p. 084006

[206] A. Taracchini *et al.*, *Effective-one-body model for black-hole binaries with generic mass ratios and spins*, Phys. Rev. D **89** (2014) (6), p. 061502

[207] B. D. Lackey, K. Kyutoku, M. Shibata, P. R. Brady and J. L. Friedman, *Extracting equation of state parameters from black hole-neutron star mergers: aligned-spin black holes and a preliminary waveform model*, Phys. Rev. D **89** (2014) (4), p. 043009

[208] S. Bernuzzi, A. Nagar, T. Dietrich and T. Damour, *Modeling the Dynamics of Tidally Interacting Binary Neutron Stars up to the Merger*, Phys. Rev. Lett. **114** (2015) (16), p. 161103

[209] T. Hinderer *et al.*, *Effects of neutron-star dynamic tides on gravitational waveforms within the effective-one-body approach*, Phys. Rev. Lett. **116** (2016) (18), p. 181101

[210] E. A. Huerta, P. Kumar, B. Agarwal, D. George, H.-Y. Schive, H. P. Pfeiffer, T. Chu, M. Boyle, D. A. Hemberger, L. E. Kidder, M. A. Scheel and B. Szilagyi, *A complete waveform model for compact binaries on eccentric orbits* (2016)

[211] L. Barack, *Gravitational self force in extreme mass-ratio inspirals*, Class. Quant. Grav. **26** (2009), p. 213001

[212] E. Poisson, A. Pound and I. Vega, *The Motion of point particles in curved spacetime*, Living Rev. Rel. **14** (2011), p. 7

[213] I. Mandel, D. A. Brown, J. R. Gair and M. C. Miller, *Rates and Characteristics of Intermediate-Mass-Ratio Inspirals Detectable by Advanced LIGO*, Astrophys. J. **681** (2008), pp. 1431–1447

[214] P. Amaro-Seoane *et al.*, *Low-frequency gravitational-wave science with eLISA/NGO*, Class. Quant. Grav. **29** (2012), p. 124016

[215] P. Amaro-Seoane *et al.*, *eLISA/NGO: Astrophysics and cosmology in the gravitational-wave millihertz regime*, GW Notes **6** (2013), pp. 4–110

[216] A. Klein *et al.*, *Science with the space-based interferometer eLISA: Supermassive black hole binaries*, Phys. Rev. **D93** (2016) (2), p. 024003

[217] L. Barack and C. Cutler, *Using LISA EMRI sources to test off-Kerr deviations in the geometry of massive black holes*, Phys. Rev. **D75** (2007), p. 042003

[218] F. D. Ryan, *Accuracy of estimating the multipole moments of a massive body from the gravitational waves of a binary inspiral*, Phys. Rev. **D56** (1997), pp. 1845–1855

[219] S. A. Hughes and R. D. Blandford, *Black hole mass and spin coevolution by mergers*, Astrophys. J. **585** (2003), pp. L101–L104

[220] C. L. MacLeod and C. J. Hogan, *Precision of Hubble constant derived using black hole binary absolute distances and statistical redshift information*, Phys. Rev. **D77** (2008), p. 043512

[221] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein and A. Petiteau, *Science with the space-based interferometer eLISA. III: Probing the expansion of the Universe using gravitational wave standard sirens*, JCAP **1604** (2016) (04), p. 002

[222] T. Hinderer and E. E. Flanagan, *Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime: Orbital motion*, Phys. Rev. D **78** (2008), p. 064028,
URL: http://link.aps.org/doi/10.1103/PhysRevD.78.064028

[223] E. E. Flanagan and T. Hinderer, *Transient resonances in the inspirals of point particles into black holes*, Phys. Rev. Lett. **109** (2012), p. 071102

[224] L. Barack and N. Sago, *Gravitational self-force on a particle in eccentric orbit around a Schwarzschild black hole*, Phys. Rev. D **81** (2010), p. 084021

[225] S. Akcay, N. Warburton and L. Barack, *Frequency-domain algorithm for the Lorenz-gauge gravitational self-force*, Phys. Rev. D **88** (2013) (10), p. 104009

[226] B. Wardell, C. R. Galley, A. Zenginoglu, M. Casals, S. R. Dolan *et al.*, *Self-force via Green functions and worldline integration*, Phys. Rev. D **89** (2014), p. 084021

[227] T. Osburn, E. Forseth, C. R. Evans and S. Hopper, *Lorenz gauge gravitational self-force calculations of eccentric binaries using a frequency domain procedure*, Phys. Rev. **D90** (2014) (10), p. 104031

[228] B. Wardell, *Self-force: Computational Strategies*, Fund. Theor. Phys. **179** (2015), pp. 487–522

[229] M. van de Meent, *Gravitational self-force on eccentric equatorial orbits around a Kerr black hole*, Phys. Rev. **D94** (2016) (4), p. 044034

[230] S. L. Detweiler, *A consequence of the gravitational self-force for circular orbits of the Schwarzschild geometry*, Phys. Rev. D **77** (2008), p. 124026

[231] L. Barack and N. Sago, *Beyond the geodesic approximation: conservative effects of the gravitational self-force in eccentric orbits around a Schwarzschild black hole*, Phys. Rev. D **83** (2011), p. 084023

[232] S. R. Dolan, P. Nolan, A. C. Ottewill, N. Warburton and B. Wardell, *Tidal invariants for compact binaries on quasi-circular orbits* (2014)

[233] S. Akcay, D. Dempsey and S. Dolan, *Spin-orbit precession for eccentric black hole binaries at first order in the mass ratio* (2016)

[234] M. van de Meent, *Self-force corrections to the periapsis advance around a spinning black hole* (2016)

[235] N. K. Johnson-McDaniel, A. G. Shah and B. F. Whiting, *Experimental mathematics meets gravitational self-force*, Phys. Rev. **D92** (2015) (4), p. 044007

[236] L. Blanchet, S. L. Detweiler, A. Le Tiec and B. F. Whiting, *High-Order Post-Newtonian Fit of the Gravitational Self-Force for Circular Orbits in the Schwarzschild Geometry*, Phys. Rev. D **81** (2010), p. 084033

[237] C. Kavanagh, A. C. Ottewill and B. Wardell, *Analytical high-order post-Newtonian expansions for spinning extreme mass ratio binaries*, Phys. Rev. **D93** (2016) (12), p. 124038

[238] A. G. Shah and A. Pound, *Linear-in-mass-ratio contribution to spin precession and tidal invariants in Schwarzschild spacetime at very high post-Newtonian order*, Phys. Rev. **D91** (2015) (12), p. 124022

[239] E. Forseth, C. R. Evans and S. Hopper, *Eccentric-orbit extreme-mass-ratio inspiral gravitational wave energy fluxes to 7PN order*, Phys. Rev. **D93** (2016) (6), p. 064058

[240] D. Bini, T. Damour and A. Geralico, *High post-Newtonian order gravitational self-force analytical results for eccentric equatorial orbits around a Kerr black hole*, Phys. Rev. **D93** (2016) (12), p. 124058

[241] T. Damour, P. Jaranowski and G. SchÃďfer, *Conservative dynamics of two-body systems at the fourth post-Newtonian approximation of general relativity*, Phys. Rev. D **93** (2016) (8), p. 084014

[242] A. Zimmerman, A. G. M. Lewis and H. P. Pfeiffer, *Redshift factor and the first law of binary black hole mechanics in numerical simulations* (2016)

[243] T. Damour, *Gravitational Self Force in a Schwarzschild Background and the Effective One Body Formalism*, Phys. Rev. D **81** (2010), p. 024017

[244] L. Barack, T. Damour and N. Sago, *Precession effect of the gravitational self-force in a Schwarzschild spacetime and the effective one-body formalism*, Phys. Rev. D **82** (2010), p. 084036

[245] D. Bini and T. Damour, *High-order post-Newtonian contributions to the two-body gravitational interaction potential from analytical gravitational self-force calculations* (2013)

[246] S. Akcay and M. van de Meent, *Numerical computation of the effective-one-body potential q using self-force results*, Phys. Rev. D **93** (2016) (6), p. 064063

[247] S. Detweiler, *Gravitational radiation reaction and second-order perturbation theory*, Phys. Rev. D **85** (2012), p. 044048,
URL: http://link.aps.org/doi/10.1103/PhysRevD.85.044048

[248] A. Pound, *Second-order gravitational self-force*, Phys. Rev. Lett. **109** (2012), p. 051101

[249] S. E. Gralla, *Second Order Gravitational Self Force*, Phys. Rev. **D85** (2012), p. 124011

[250] A. Pound and J. Miller, *Practical, covariant puncture for second-order self-force calculations*, Phys. Rev. **D89** (2014) (10), p. 104020

[251] B. Wardell and N. Warburton, *Applying the effective-source approach to frequency-domain self-force calculations: Lorenz-gauge gravitational perturbations*, Phys. Rev. D **92** (2015) (8), p. 084019

[252] J. Miller, B. Wardell and A. Pound, *Second-order perturbation theory: the problem of infinite mode coupling* (2016)

[253] A. Pound, *Second-order self-force: a progress report*, talk given at GR21, Columbia University, New York, USA, July 2016. Slides available at http://www.gr21.org/files/Adam_Pound_-_Pound_(1).pdf.gz

[254] Y. M. M. S. T. Tanaka, *Gravitational radiation reaction to a point particle*, Phys. Rev. D (1996), p. 55.3457

[255] T. C. Quinn and R. M. Wald, *Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime*, Phys. Rev. D **56** (1997), pp. 3381–3394, URL: http://link.aps.org/doi/10.1103/PhysRevD.56.3381

[256] A. Pound, *Self-consistent gravitational self-force*, Phys. Rev. **D81** (2010), p. 024023

[257] A. Pound, *Nonlinear gravitational self-force: Field outside a small body*, Phys. Rev. D **86** (2012), p. 084019

[258] S. R. Dolan and L. Barack, *Self-force via m-mode regularization and 2+1D evolution: III. Gravitational field on Schwarzschild spacetime*, Phys. Rev. D **87** (2013), p. 084066

[259] P. Diener, I. Vega, B. Wardell and S. Detweiler, *Self-consistent orbital evolution of a particle around a Schwarzschild black hole*, Phys. Rev. Lett. **108** (2012), p. 191102

[260] A. Pound and E. Poisson, *Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals*, Phys. Rev. **D77** (2008), p. 044013

[261] J. R. Gair, E. E. Flanagan, S. Drasco, T. Hinderer and S. Babak, *Forced motion near black holes*, Phys. Rev. D **83** (2011), p. 044037

[262] N. Warburton, S. Akcay, L. Barack, J. R. Gair and N. Sago, *Evolution of inspiral orbits around a Schwarzschild black hole*, Phys. Rev. D **85** (2012), p. 061501

[263] T. Osburn, N. Warburton and C. R. Evans, *Highly eccentric inspirals into a black hole*, Phys. Rev. **D93** (2016) (6), p. 064024

[264] Y. Mino, *Perturbative approach to an orbital evolution around a supermassive black hole*, Phys. Rev. D **67** (2003), p. 084027

[265] N. Sago, T. Tanaka, W. Hikida, K. Ganz and H. Nakano, *The Adiabatic evolution of orbital parameters in the Kerr spacetime*, Prog. Theor. Phys. **115** (2006), pp. 873–907

[266] S. Drasco, E. E. Flanagan and S. A. Hughes, *Computing inspirals in Kerr in the adiabatic regime. I. The Scalar case*, Class. Quant. Grav. **22** (2005), pp. S801–846

[267] R. Fujita, W. Hikida and H. Tagoshi, *An Efficient Numerical Method for Computing Gravitational Waves Induced by a Particle Moving on Eccentric Inclined Orbits around a Kerr Black Hole*, Prog. Theor. Phys. **121** (2009), pp. 843–874

[268] S. A. Hughes, *Adiabatic and post-adiabatic approaches to extreme mass ratio inspiral* (2016), URL: https://inspirehep.net/record/1414738/files/arXiv:1601.02042.pdf

[269] S. A. Hughes, *Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational wave emission. II. Inspiral trajectories and gravitational wave forms*, Phys. Rev. D **64** (2001), p. 064004, [Erratum: Phys. Rev.D88,no.10,109902(2013)]

[270] K. A. Lackeos and L. M. Burko, *Self-forced gravitational waveforms for Extreme and Intermediate mass ratio inspirals*, Phys. Rev. **D86** (2012), p. 084055

[271] A. Taracchini, A. Buonanno, G. Khanna and S. A. Hughes, *Small mass plunging into a Kerr black hole: Anatomy of the inspiral-merger-ringdown waveforms*, Phys. Rev. D **90** (2014) (8), p. 084025

[272] A. Pound, *Motion of small bodies in curved spacetimes: an introduction to gravitational self-force* (Springer, 2015), vol. 179 of *Fundamental Theories of Physics*

[273] Y. Mino, *Adiabatic expansion for metric perturbation and the condition to solve the gauge problem for gravitational radiation reaction problem*, Prog. Theor. Phys. **115** (2006), pp. 43–61

[274] A. Pound and E. Poisson, *Multi-scale analysis of the electromagnetic self-force in a weak gravitational field*, Phys. Rev. D **77** (2008), p. 044012

[275] Y. Mino and R. Price, *Two-timescale adiabatic expansion of a scalar field model*, Phys. Rev. D **77** (2008), p. 064001

[276] A. Pound, *Singular perturbation techniques in the gravitational self-force problem*, Phys. Rev. D **81** (2010), p. 124009

[277] V. Balasubramanian, A. Buchel, S. R. Green, L. Lehner and S. L. Liebling, *Holographic Thermalization, Stability of AntiâĂŞde Sitter Space, and the Fermi-Pasta-Ulam Paradox*, Phys. Rev. Lett. **113** (2014) (7), p. 071601

[278] A. Pound, *Second-order perturbation theory: problems on large scales*, Phys. Rev. **D92** (2015) (10), p. 104047

[279] U. Ruangsri and S. A. Hughes, *Census of transient orbital resonances encountered during binary inspiral*, Phys. Rev. **D89** (2014) (8), p. 084036

[280] J. Brink, M. Geyer and T. Hinderer, *Orbital resonances around Black holes*, Phys. Rev. Lett. **114** (2015) (8), p. 081102

[281] S. Isoyama, R. Fujita, H. Nakano, N. Sago and T. Tanaka, *Evolution of the Carter constant for resonant inspirals into a Kerr black hole: I. The scalar case*, PTEP **2013** (2013) (6), p. 063E01

[282] P. D. D'Eath, *Dynamics of a small black hole in a background universe*, Phys. Rev. **D11** (1975), pp. 1387–1403

[283] K. S. Thorne and J. B. Hartle, *Laws of motion and precession for black holes and other bodies*, Phys. Rev. **D31** (1984), pp. 1815–1837

[284] S. L. Detweiler, *Radiation reaction and the selfforce for a point mass in general relativity*, Phys. Rev. Lett. **86** (2001), pp. 1931–1934

[285] S. E. Gralla and R. M. Wald, *A Rigorous Derivation of Gravitational Self-force*, Class. Quant. Grav. **25** (2008), p. 205009, [Erratum: Class. Quant. Grav.28,159501(2011)]

[286] R. W. S. Gralla, A. Harte, *A Rigorous Derivation of the Electromagnetic Self Force* (2009)

[287] A. Pound, *Nonlinear gravitational self-force: second-order equation of motion* (2017)

[288] R. M. Wald, *General Relativity* (Chicago Univ. Pr., Chicago, USA, 1984)

[289] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973)

[290] S. Isoyama, A. Pound, T. Tanaka and K. Yamada, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. IV. Dynamics of the near-horizon region* ((In prep.))

[291] E. E. Flanagan, T. Hinderer, J. Moxon and A. Pound, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. III. Dynamics of the far zone waves* ((In prep.))

[292] G. O. A. Team, *The ESA-L3 Gravitational Wave Mission* (2016)

[293] J. Kevorkian and J. D. Cole, *Multiple Scale and Singular Perturbation Methods* (Springer, New York, 1996)

[294] M. Bruni, S. Matarrese, S. Mollerach and S. Sonego, *Perturbations of space-time: Gauge transformations and gauge invariance at second order and beyond*, Class. Quant. Grav. **14** (1997), pp. 2585–2606

[295] J. Vines and Ã. Ã. Flanagan, *Is motion under the conservative self-force in black hole space-times an integrable Hamiltonian system?*, Phys. Rev. **D92** (2015), p. 064039

[296] M. Van De Meent and N. Warburton, *Fast Self-forced Inspirals* (2018)

[297] L. Barack and A. Ori, *Gravitational selfforce and gauge transformations*, Phys. Rev. **D64** (2001), p. 124003

[298] E. Rosenthal, *Second-order gravitational self-force*, Phys. Rev. D **74** (2006), p. 084018, URL: http://link.aps.org/doi/10.1103/PhysRevD.74.084018

[299] V. I. Arnold, *Mathematical methods of classical mechanics* (Springer, 1995)

[300] E. Fiorani, G. Giachetta and G. Sardanashvily, *LETTER TO THE EDITOR: The Liouville-Arnold-Nekhoroshev theorem for non-compact invariant manifolds"*, Journal of Physics **36** (2003)

[301] L. Blanchet and T. Damour, *Hereditary effects in gravitational radiation*, Phys. Rev. **D46** (1992), pp. 4304–4319

[302] L. Barack, *Late time dynamics of scalar perturbations outside black holes. 2. Schwarzschild geometry*, Phys. Rev. **D59** (1999), p. 044017

[303] A. Pound, *Gauge and motion in perturbation theory*, Phys. Rev. **D92** (2015) (4), p. 044021

Chapter 5

The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions, II. Dynamics of the strong-field region

Coauthors:

Éanna Flanagan, Cornell University

Tanja Hinderer, University of Maryland

Adam Pound, University of Southampton

5.1 Introduction and motivation

Self force methods are currently the most efficient and most accurate method for computing a waveform from high mass ratio binaries in the strong-field regime, including both the EMRIs that will be detectable by LISA [304] and intermediate mass ratio inspirals (IMRIs) of mass ratio $\mu/M \sim 100$ that could be detected soon by LIGO/Virgo [305]. The waveforms required to make best use of LISA data must track the phase of the waveform to an accuracy $\ll 1$ radian [306, 307]. In the self force expansion away from resonance, the phase value can be expanded using $\varepsilon \equiv \mu/M$ as $\phi(\varepsilon) = \varepsilon^{-1}\phi^{(0)} + \phi^{(1)} + \varepsilon\phi^{(2)}$, so the phase error requirement demands solutions through $\phi^{(1)}$ (post-adiabatic order), which requires computation of dissipative second order and all first order effects [306]. The long-lived inspiral of EMRIs ($\mathcal{O}(M/\mu)$ orbits) implies that the sub-radian accuracy goal places stringent requirements on the accuracy of the inspiral evolution. There are three principal

stages to a self force computation, each of which must be achieved with sufficient accuracy to satisfy the ultimate goal of sub-radian waveform accuracy. First, the self force itself is computed for a given stage of the inspiral from the analytically derived forms in terms of the regularized metric perturbations. The self force must then be used to evolve the orbit of the small companion, appropriately back-reacting the effect of the motion on the sourced metric perturbation. Finally, a waveform can be computed from the motion and the sourced metric perturbation. The self force methods have proved extraordinarily robust, and methods based on self force expansions can now evolve generic orbits in Schwarzschild [308–312] and Kerr [313] backgrounds at first order. The self force techniques have provided extremely accurate results [314], and computed a wide variety of orbit effects [315–319] in the strong field region.

The goal in this series of papers is to construct a comprehensive scheme to achieve the sub-radian accuracy goal for a generic class of orbits. We propose the multiscale framework as an efficient, practical method for all components of a self-force computation which can capture all slowly varying effects of the binary and reach sub-radian accuracy. Many of the techniques we use to assemble the comprehensive algorithm are built on previously developed techniques of the self force community, extended and adapted to the multiscale framework. We review below many of the salient details of the techniques used to compute self force waveforms, after which we describe the multiscale expansion and discuss its relation to prior methods.

The direct evaluation of the self force of an orbit requires regularization of metric perturbation, which becomes singular near the worldline of the small companion. Given an approximate orbit, and the metric perturbation in the vicinity of that orbit, there are a number of robust methods that have been developed of the evaluation of the instantaneous self force [312]. The mode-sum regularization approach [320, 321] relies on a series of analytically derived spherical harmonic mode coefficients for the singular part of the metric perturbation. In a spectral computation of the metric perturbation, these analytically computed singular modes may then individually be subtracted from the computed modes of the spacetime to develop a regular metric perturbation which can be used in the self force. Another frequently used method is the effective source method [322, 323], which is more generally applicable, and explicitly subtracts a source term analytically derived for the singular part of the metric perturbation. The metric perturbation that is then computed from the

wave equation is a regularized metric perturbation, rather than the physical metric perturbation. Most recently, methods of computing a regularized self field directly from regular components of the Green's function bitensor, initially computed over the computational domain, then convolved with the worldline to obtain an explicit computation of the formal expressions for the regular self field [312].

To obtain a computation of the self accelerated worldline, the self force must be used to evolve the non-geodesic worldline. The self force community has developed a number of methods for computing the orbit of the small companion, each of which strikes a different trade-off between accuracy of the inspiral and speed of the computation. The fastest and coarsest method that has been developed is referred to as the adiabatic approximation. In the adiabatic approximation [324–328], all effects of the self force that do not result in dissipation of conserved orbit parameters are neglected, and the orbit is evolved entirely via the computation of first-order fluxes. A significantly more thorough technique is the method of osculating geodesics [329, 330], implemented in [331, 332], in which the full self force is computed from the instantaneous tangent geodesic at each moment of the inspiral. The osculating geodesics method loses accuracy, as the self force is computed as though the small companion eternally occupies the instantaneous geodesic, so approximates poorly the contribution to the metric perturbation sourced significantly in its past. Currently, the most accurate but most costly method is the self-consistent method [333], for which implementations include [334,335]. The self consistent evolution makes no dynamical approximations beyond numerical gridding and the expansion of the mass ratio. The orbit of the small companion is evolved simultaneously with the metric perturbation it sources, and the self force is computed at each time step.

Finally, a self force computation must translate the computed dynamics of the worldline and strong field metric perturbation to a resulting gravitational waveform. As in the case of the worldline computation, there is a significant trade-off between computational speed and accuracy of the waveform. The fastest and least accurate method is known as the semi-relativistic method [336,337], which estimates the radiation from a binary source by approximating the source as a pair of test masses in Minkowski spacetime and finding the radiation via quadrupole radiation calculation. Another method, which is analogous to the osculating geodesics method of source evolution, is the snapshot evolution of waveforms [338]. The snapshot evolution evaluates the waveform sourced by

several geodesic sources by carefully solving the black hole perturbation equations, then stitches together several waveforms as the binary traces through the parameter space of geodesics. The most precise currently applied method [339], and the most computationally expensive, is one which solves for the full self consistent waveform from the evolution of the worldline, either starting from the worldline itself, or by propagating out to future null infinity $\mathcal{I}^+$ the metric perturbation used for the self force.

We propose the multiscale approximation method as an algorithm which will both recover the waveform at the sub-radian accuracy goal and incorporate any efficiency gains which are possible from the near-periodic orbital dynamics. The multiscale approximation method is a well-documented technique for solving weakly nonlinear oscillator and wave equations which rigorously separates the dynamics into periodic behavior and slow nonlinear reaction behavior. These mathematical techniques are well-suited to the scenario of an EMRI, in which the orbit evolves near-periodically on the timescale M, and slowly loses energy on timescale μ. A full multiscale approximation should permit many of the optimizations associated with osculating geodesics and snapshot waveform evolution without sacrificing any phase accuracy, by identifying the portions of the waveform calculation which can be consistently evolved through the phase and geodesic variables. In addition, the multiscale approximation will be able to incorporate the optimizations associated with the near-identity transformation of orbits, which have recently been shown to give significant computational speedup [340].

Multiscale approximation methods were first suggested as a useful method for expanding the orbits of high mass ratios in [341, 342]. In previous explorations, we have derived the multiscale dynamics of generic orbits in Kerr away from resonances [306], and performed a multiscale back-reaction computation of a representative scalar model, which demonstrated the need for matching regions at $|r^*| \gg M$. To offer a complete technique for evaluating self-forced inspirals to sub-radian accuracy, we require a full generalization of [343] to the multiscale expansion of the EFE evolved together with the worldline equations found in [306]. The results from [306] are two sets of equations for the action and action-angle variables, each of which with corrections from the self force determined by metric perturbation. We review in detail the derived orbital equations in Section 5.2.2. In the previous paper in this series [344], we give the overview of our multiscale algorithm for

the computation of the metric perturbation, and present the equations necessary for evolving the orbit at adiabatic order.

In this paper, we develop the full post-adiabatic multiscale formalism for the strong-field region $|r^*| \ll M/\varepsilon$, which we refer to as the interaction zone. The dynamics of the interaction zone involve both the multiscale development of the nonlinear wave equation, and the backreaction of the self force on the multiscale orbital motion through second order in the mass ratio. In Section 5.2, we give a brief survey of the results from previous papers in the series, with emphasis on the field equations which will be used heavily in the remainder of this paper. We then give the treatment of the second order wave equation in the multiscale formalism, including the extraction of the slowly varying mass and spin in the Teukolsky-Lousto-Campanelli wave equations in section 5.3. Finally, we present in section 5.4 a second-order extension to the flux balance laws for deriving the energy, angular momentum, and Carter constant evolution of the orbit from asymptotic fluxes of second order quantities.

5.2 A overview of multiscale formalism

5.2.1 Separation of scales and approximation regions

We define the mass ratio $\varepsilon = \mu/M \ll 1$, and introduce a one-parameter family of spacetimes, described by a manifold and an ε-dependent metric and stress energy tensor,

$$(\mathcal{M}, g_{\mu\nu}(\varepsilon), T_{\mu\nu}(\varepsilon)), \tag{5.1}$$

which satisfies the Einstein field equation,

$$G_{\mu\nu}[g(\varepsilon)] = 8\pi T_{\mu\nu}(\varepsilon), \tag{5.2}$$

such that the metric $g_{\mu\nu}$ and stress energy tensor $T_{\mu\nu}$ are jointly smooth in the spacetime coordinates and ε for all $\varepsilon > 0$. We refer the interested reader to the more careful introduction in Chapter 4 for the remaining mathematical preliminaries. For the purposes of the current paper, it is sufficient to note the one-parameter family, and the set of specific approximations we use in the various regions of the spacetime, which follow.

The multscale approximation scheme assembled in Chapter 4 introduces several regions, each connected by a matched asymptotic expansion. Our tapestry of approximations is required because no single one of our approximations is convergent through second order over the full spacetime. We use two spatial scales to determine the scales of the various regions: the distance from the small companion $\tilde{r}$, and the radial tortoise coordinate of the background spacetime r^*. In terms of these variables, our four principal regions and their respective approximation schemes are:

- The *body zone* or *puncture zone*, where $\tilde{r} \ll M$. This region is evaluated using techniques from the self consistent expansion, and in the full resulting algorithm only the punctures from this region appear.

- The *interaction zone*, where $\tilde{r} \gg \mu$ and $|r^*| \ll M/\varepsilon$. This paper focuses on the dynamics of the interaction zone, which involves a multiscale expansion of the field equations and takes advantage of results from self consistent formalism for the backreaction of the worldline source.

- The *far zone*, where $r^* \gg M$ describes positions far from the entire inspiral system. In this region, we use an adapted geometric optics formalism, with a Post-Minkowski expansion used to determine quasistatic contributions. Full details of the second order computation in this region can be found in the forthcoming paper [345]. An overview of the technues used in the far zone to post-adiabatic order can be found in Chapter 6

- The *near horizon zone*, where $-r^* \gg M$ describes positions very near the massive black hole horizon. In this region, we use an expansion adapted to the near-horizon metric. Full details of the treatment of the multiscale formalism applied to this region can be found in the forthcoming paper [346].

Our core approximation, and the one used in the interaction zone is the multiscale approximation. In this method, we introduce a pair of time variables $\{q^A, \tilde{w}\}$, for which all physical observables depend only periodically on q^A, and $\tilde{w}$ parameterizes the slow radiation-reaction effects. However, as discussed in detail in Section 4.4, the multiscale approximation fails at distances similar to the radiation reaction time $r \sim M/\varepsilon$, as the finite speed of propagation is poorly approximated on such distances. The long-distance failure of the multiscale approximation was explored thoroughly in a

scalar analog model in [343], in which it was demonstrated that the problems at large distances cause the second order multiscale approximation to diverge when used over the full spacetime.

The multiscale procedure consists of a function ansatz over an expanded domain, along with a map from that larger domain to the physical domain. Specifically, we use the pair of time variables $\{q^A, \tilde{w}\}$, such that the map to the physical spacetime is given by functions,

$$\tilde{w} = \varepsilon w, \tag{5.3a}$$

$$\frac{dq^A}{dw} = \Omega^{(0)A}(\tilde{w}) + \varepsilon\Omega^{(1)A}(\tilde{w}) + \mathcal{O}(\varepsilon^2), \tag{5.3b}$$

where we use physical time variable w, related to Boyer-Lindquist time t by

$$w = t + h(r), \tag{5.4}$$

for generic height function $h(r)$, such that $h(r)$ approaches r at $r^* \gg M$ and approaches $-r$ at $-r^* \gg M$. These conditions ensure that surfaces of constant w are asymptotically null. The asymptotically null choice of time variable will improve the conditions required for matching the interaction zone to the Far and Near-Horizon Zones. Full details of those matching computations will be given in papers [345] and [346] for the Far and Near-Horizon matching, respectively. Using this modified time variable, the background metric in the modified Boyer-Lindquist coordinates is,

$$\begin{aligned} ds^2 = &-\left(1 - \frac{2Mr}{\Sigma}\right) dw^2 - \frac{4aMr\sin^2\theta}{\Sigma} dw d\phi + (\varpi^4 - \Delta a^2 \sin^2\theta)\frac{\sin^2\theta}{\Sigma} d\phi^2 + \Sigma d\theta^2 \\ &+ \left(\frac{\Sigma}{\Delta} - h'(r)^2\left(1 - \frac{2Mr}{\Sigma}\right)\right) dr^2 + h'(r)\frac{4aMr\sin^2\theta}{\Sigma} dr d\phi - 2\left(1 - \frac{2Mr}{\Sigma}\right) h'(r) dr dw. \end{aligned} \tag{5.5}$$

The fast time variables q^A are those associated with the action-angle formalism described in Section 5.2.2, and the frequencies in (5.3b) are determined with respect to physical coordinate time variable w. Therefore, the map $\{x^i, w\} \to \{x^i, q^A, \tilde{w}\}$ off the worldline has the interpretation of parameterizing physical quantities by the phase coordinate value equal to the value of q^A of the orbit on the slice of constant w which is shared with the point $\{x^i, w\}$.

In the Interaction Zone, the multiscale approximation is developed by constructing metric and

worldline functions over the expanded time domain $\{q^A, \tilde{w}\}$:

$$g_{\mu\nu}(x^i, q^A, \tilde{w}, \varepsilon) = g^{(0)}_{\mu\nu}(x^i) + \varepsilon h^{(1)}_{\mu\nu}(x^i, q^A, \tilde{w}) + \varepsilon^2 h^{(2)}_{\mu\nu}(x^i, q^A, \tilde{w} + \mathcal{O}(\varepsilon^3), \tag{5.6a}$$

$$z^\mu(q^A, \tilde{w}, \varepsilon) = z^{(0)\mu}(q^A, \tilde{w}) + \varepsilon z^{(1)\mu}(q^A, \tilde{w}) + \varepsilon^2 z^{(2)\mu}(q^A, \tilde{w}) + \mathcal{O}(\varepsilon^3). \tag{5.6b}$$

The perturbative expansion 5.6 is robust in the high mass ratio limit, provided we demand that the system does not enter an orbital resonance [307, 347, 348]. Orbital resonances in general introduce a third timescale of order $\mathcal{O}(\sqrt{\varepsilon}M)$ which describes the timescale for which the resonant effects are dominant. In the current work, we assume that the system does not enter any resonance during the multiscale evolution. Practically, circular inclined orbits and equatorial orbits possess no resonances due to the explicit elimination of one of the relevant angle variables. For generic inclined orbits in Kerr, it has been shown that [349] orbits typically pass through at least one low-order resonance over the full evolution. Inclusion of resonances in the multiscale formalism will be the subject of future work.

In the following two subsections, we will review first the overview of the computational method for the multiscale orbit equations (Section 5.2.2), followed by the basic expansions of the multiscale metric perturbation (Section 5.2.3).

5.2.2 | Orbital equations through second order

The material in this section is a highly abridged summary of the multiscale orbit derivations in [306], and the incorporation into the multiscale formalism discussed in [344]. Those computations construct a phase space parameterized by the generalized action-angle variables (q_α, J_α), and the geodesic parameters $P^M = \{E, L_z, Q\}$ depend exclusively on the action variables $P_\alpha = P_\alpha(J_\beta)$. The equations of motion for the geodesic Hamiltonian system can be written,

$$\frac{dq_A}{dw} = \frac{\partial H(J)}{\partial J^A} \equiv \omega_A(J) \tag{5.7a}$$

$$\frac{dJ_A}{dw} = -\frac{\partial H(J)}{\partial q^A}. \tag{5.7b}$$

The Hamiltonian system is then promoted to a dissipative system via generic methods explained in [350]. The result of applying that formalism to the near-Hamiltonian system of self forced orbits

in Kerr gives the coupled equations of motion [306, 344],

$$\frac{dq^A}{dw} = \omega^A(P^M) + \varepsilon f^A{}_\mu a^{(1)\mu}(q^A, P^M, h^{(1)}) + \varepsilon^2 f^A{}_\mu a^{(2)\mu}(q^A, P^M, h^{(2)}, h^{(1)}) + \mathcal{O}(\varepsilon^3), \tag{5.8a}$$

$$\frac{dP^M}{dw} = \varepsilon F^M{}_\mu a^{(1)\mu}(q^A, P^M, h^{(1)}) + \varepsilon^2 F^M{}_\mu a^{(2)\mu}(q^A, P^M, h^{(2)}, h^{(1)}) + \mathcal{O}(\varepsilon^3), \tag{5.8b}$$

where $a^{(n)\mu}$ is the order-n self acceleration derived from [351, 352], and the functions $f^A{}_\mu$ and $F^M{}_\mu$ are calculated in [306] as

$$f^A{}_\mu = \frac{\partial q^A}{\partial p^\mu}, \tag{5.9a}$$

$$F^M{}_\mu = \frac{\partial J^M}{\partial P^N}\frac{\partial P^N}{\partial p^\mu}. \tag{5.9b}$$

The momenta in these functions are the caoninical momenta of the Kerr geodesic Hamiltonian. The remaining details of the celestial mechanics functions and the explicit values for the force transformation functions $f^A{}_\mu$ and $F^M{}_\mu$ are detailed in [306].

The self-acceleration which appears in the equations of motion (5.8) is determined entirely from the state of the orbit and the metric perturbation. However, the backreaction of the orbit evolution to the metric perturbation requires an expression for the worldline itself as a function of the action angle variables and the geodesic parameters, $z^\mu(q^A, P^M)$. The procedure for determining those functions is to perturbatively expand the worldline equations of motion in terms of the momenta, and substituting the geodesic parameters as functions of the fast and slow time variables. The full procedure for deriving the leading and subleading worldline functions $z^{(1)\mu}(q^A, P^M)$, and the results thereof are detailed in Appendix 5.A. The corrected worldline is then used to derive the correction to the metric perturbation near the worldline.

5.2.3 Overview of the multiscale procedure for the Einstein field equation

To extract a reliable waveform from an efficient and accurate multiscale computation of the orbit from Section 5.2.2 and [306], we must peform a similar expansion for the field equations themselves [344]. The motiviation for the multiscale expansion of the spacetime metric follows directly from the multiscale expansion of the worldline. Given that the $\mathcal{O}(\mu)$ stress energy source has only periodic time dependence, then the time-translation symmetry of the background spacetime ensures that the perturbations are similarly periodic, provided resonances are neglected.

The multiscale expansion for the metric perturbation takes the form (5.6). It should be understood that (5.6) is a function over a significantly larger domain than the ordinary spacetime degrees of freedom. Similar to the orbit parameters, the metric perturbations depend on four time variables $\{\tilde{w}, q^A\}$. The metric perturbation as a function of physical time variables is obtained by evaluating the multiscale functions at the physical times determined by the frequencies,

$$h^{(n)}_{\mu\nu}(x^\mu) = h^{(n)}_{\mu\nu}(\tilde{w}, q^A, x^\mu)\Big|_{\tilde{w}=\varepsilon w, q^A = q^A(w)}, \tag{5.10}$$

where $q^A(w)$ is computed from (5.8a).

In the following discussion, we use superscripts of the covariant derivatives to denote the order of ε dependence from the time derivatives. The subleading covariant derivative has up to one slow time derivative and expands the subleading frequency dependence associated with fast time dependence. Explicitly,

$$\nabla^{(0)}_\mu = \nabla_\mu|_{\partial_w \to \Omega^{(0)A}\partial_A}, \tag{5.11a}$$

$$\nabla^{(1)}_\mu = w_\mu(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_A), \tag{5.11b}$$

$$\nabla^{(2)}_\mu = w_\mu \Omega^{(2)A}\partial_A, \tag{5.11c}$$

where $\tilde{w}$ describes the slow time, and w_μ is the vector which generates translations in w. The superscript of the relaxed Einstein field operator $E^{(n)}_{\mu\nu}$ and the Ricci tensor $R^{(n)}_{\mu\nu}$ similarly denote the order of expansion of the time derivatives, where their arguments imply the expansion in powers of the metric perturbations. As discussed in detail in [344] Section II D, if the expanded multiscale ansatz (5.6) obeys a multiscale version of a given wave equation obtained by expanding derivatives order-by-order according to (5.11), then the physically projected metric perturbation (5.10) obeys the corresponding physical wave equation.

Similar to the orbit equations of motion, the effect of the multiscale expansion on the field equations can be determined by using the time derivative expansions (5.11). The conclusion is that, for every order above the first, there arises a source correction associated with the slow variation of parameters at lower orders. Schematically, the perturbative Einstein field equations take the form,

$$R^{(0)}_{\mu\nu}[h^{(1)}] = 8\pi \bar{T}_{\mu\nu}, \tag{5.12a}$$

$$R^{(0)}_{\mu\nu}[h^{(2)}] = -R^{(1)}_{\mu\nu}[h^{(1)}] - R^{(0)}_{\mu\nu}[h^{(1)}, h^{(1)}]. \tag{5.12b}$$

In addition to the source correction, the multiscale expansion requires an additional step which computes the slow time dependence of each order in the expansion by demanding consistency with the multiscale ansatz. Much of the slow time dependence of the metric perturbation is determined by the slow time dependence of the worldline, and the only remaining unfixed parameters are those parameters of the spacetime permitted by the no-hair theorem to evolve secularly. In particular, we need only determine the slow variation of the central black hole mass $\delta M(\tilde{w})$ and spin $\delta a(\tilde{w})$.

In principal, there are many secularly evolving parameters of the spacetime we might worry about, including effects of overall inertial frame shift of the central black hole as compared to distant stars as well as the more general class of subtle conserved charges associated with asymptotic (BMS) symmetries. For post-adiabatic corrections, however, the overall shift in any of these parameters is at most $\mathcal{O}(\varepsilon)$ over the full inspiral, and can be converted to a slowly evolving $\mathcal{O}(\varepsilon^2)$ contribution via a gauge transformation. Therefore, for the purposes of the strong-field dynamics, the effect of all parameters apart from total mass and charge of the central black hole appear as quasistatic alterations to the second order metric perturbations and can be neglected for the computations presented in this paper. For the purposes of extracting waveforms, the slowly evolving overall redshift of the system from center of mass motion may need to be considered.

The first order metric perturbation is singular at the position of the small companion's worldline, $z^\mu(\tilde{w}, q^A)$, so a modification to the direct computation from the Einstein field equations (5.12) is required. For the purposes of developing our computational framework, we will contruct a smooth metric perturbation motivated by the self consistent computation method [333] and the effective source numerical techniques [323].

Formally, we expand the Einstein field equations separately at small distances $r \ll M$ to the small companion and at large distances from the small companion $r \gg m$. In the region $r \ll M$, we must explicitly subtract off a suitable singular component of the metric perturbation; otherwise, the singularity from the quadratic source causes the second order wave equation to be ill-defined [353]. We denote the subtracted singular field $h^{\mathcal{P}}$, as the singular field that will be used will likely be

similar to the puncture metric derived in [353] for realistic computations.

$$R^{(0)}_{\mu\nu}[h^{(1)\mathcal{R}}] = -R_{\mu\nu}[h^{(1)\mathcal{P}}] + 8\pi\bar{T}_{\mu\nu} \qquad (r \ll M), \tag{5.13a}$$

$$R^{(0)}_{\mu\nu}[h^{(2)\mathcal{R}}] = -R^{(1)}_{\mu\nu}[h^{(1)}] - R^{(0)}_{\mu\nu}[h^{(1)}, h^{(1)}] - R^{(0)}_{\mu\nu}[h^{(1)\mathcal{P}}] \qquad (r \gg M), \tag{5.13b}$$

$$R^{(0)}_{\mu\nu}[h^{(1)}] = 0 \qquad (r \gg \mu), \tag{5.13c}$$

$$R^{(0)}_{\mu\nu}[h^{(2)}] = -R^{(1)}_{\mu\nu}[h^{(1)}] - R^{(0)}_{\mu\nu}[h^{(1)}, h^{(1)}] \qquad (r \gg \mu). \tag{5.13d}$$

The set of equations (5.13) can then be solved simultaneously by imposing that the internal and external metrics are consistent in the shared overlap region $\mu \ll r \ll M$. However, the information from the matching computation can, and has, been used to determine the self force equations of motion [351, 352, 354], as well as a covariant form for singular part of the metric perturbation $h^{\mathcal{P}}$ through second order [353]. The set of analytical results is then sufficient to reduce the problem to a single wave equation with an effective source determined analytically from the puncture metric $h^{\mathcal{P}}$ in terms of the homogeneous components of the metric perturbations at lower orders and the parameters of the small companion.

We then write the effective source field equations which will be used for the Interaction Zone computation,

$$R^{(0)}_{\mu\nu}[\tilde{h}^{(1)}] = -R_{\mu\nu}[h^{(1)\mathcal{P}}] + 8\pi\bar{T}_{\mu\nu} \equiv S^{\text{eff}(1)}_{\mu\nu}, \tag{5.14a}$$

$$R^{(0)}_{\mu\nu}[\tilde{h}^{(2)}] = -R^{(1)}_{\mu\nu}[h^{(1)}] - R^{(0)}_{\mu\nu}[h^{(1)}, h^{(1)}] - R^{(0)}_{\mu\nu}[h^{(1)\mathcal{P}}] \equiv S^{\text{eff}(2)}_{\mu\nu}, \tag{5.14b}$$

where the equations are now evaluated over the full spacetime. Note that the metric perturbation $\tilde{h}$ will be approximately the regular field $h^{\mathcal{R}}$ very close to the worldline, and approximately h far from the worldline. For numerical calculations, the puncture field is often taken to go sharply to zero, in which case the field $\tilde{h}$ should be identical to h outside the region of support of $h^{\mathcal{P}}$.

The first order wave equations may be solved either in the Lorenz gauge or by way of the Teukolsky wave equation. Consider first the computation which makes use of the Lorenz gauge wave equation. We will first impose a natural generalization of the Lorenz gauge condition to multiscale dependence,

$$0 = \nabla^{(0)}_{\mu}\bar{h}^{(1)\mu\nu}, \tag{5.15a}$$

$$0 = \nabla^{(0)}_{\mu}\bar{h}^{(2)\mu\nu} + \nabla^{(1)}_{\mu}\bar{h}^{(1)\mu\nu}, \tag{5.15b}$$

where $\bar{h}^{(n)\mu\nu}$ denotes the trace-reversed metric perturbation. This gauge condition implies the multiscale field equations in terms of the relaxed Einstein field operator,

$$E^{(0)}_{\mu\nu}[h^{(1)}] = 8\pi \bar{T}_{\mu\nu}, \tag{5.16a}$$

$$E^{(0)}_{\mu\nu}[h^{(2)}] = \delta^2 R^{(0)}_{\mu\nu}[h^{(1)}, h^{(1)}] - E^{(1)}_{\mu\nu}[h^{(1)}]. \tag{5.16b}$$

Applying the gauge condition (5.15) to the general field equations (5.14) then implies multiscale form for the effective source field equations,

$$E^{(0)}_{\mu\nu}[\tilde{h}^{(1)}] = 8\pi T_{\mu\nu} - E^{(0)}_{\mu\nu}[h^{(1)\mathcal{P}}], \tag{5.17a}$$

$$E^{(0)}_{\mu\nu}[\tilde{h}^{(2)}] = \delta^2 R^{(0)}_{\mu\nu}[h^{(1)}, h^{(1)}] - E^{(1)}_{\mu\nu}[h^{(1)}] - E^{(0)}_{\mu\nu}[h^{(2)\mathcal{P}}]. \tag{5.17b}$$

In this form of the gauge choice[1], the wave equation (5.17) determines all parts of the metric perturbation aside from the slow evolution of the mass $\delta M(\tilde{t})$ and spin $\delta a(\tilde{t})$. At first order, the relaxed Einstein field equation (5.17) gives a collection of ten coupled wave equations, for which the a method of computation is developed in [308].

An alternative computational strategy at first order is to make use of the Teukolsky master equation [355] to compute instead the first order Weyl scalars ψ_0 and ψ_4 (Newman-Penrose definitions relevant to the discussion of the Weyl scalar wave equations can be found in Appendix 5.C). Using Teukolsky modes often grants a significant performance improvement, due to the comparatively few degrees of freedom and the separability of the ψ_0 and ψ_4 wave equations. For computations of the self-force waveform to post-adiabatic order, however, there is the necessary cost of reconstructing the full metric perturbations from the ψ_0 and ψ_4 solutions. The full metric perturbation $h^{(1)}_{\mu\nu}$ is required both for the second-order Teukolsky-Lousto-Campanelli scalar wave equations [356] and the conservative part of the first order self force, all of which is required for post-adiabatic waveform accuracy [306, 344]. The full description of the adaptation of the scalar wave equation procedures through second order to the multiscale framework is given in Section 5.3.

There is the potential to gain significant performance improvements from avoiding the requirement of metric reconstruction at second order. In Section 5.4, we present, for the first time, a derivation of the second order flux balance laws for the energy and momentum of the worldline.

[1] see Section 4.3.6 for a discussion of alternative forms

The post-adiabatic waveform depends exclusively on the dissipative effects from the second-order self force [306], which can be extracted from a combination of information from the second-order Weyl scalars and full first-order metric.

The self force computation of the orbit evolution, either from fluxes or directly from the self force expressions applied to the orbit equations from Section 5.2.2 and [306], determines all of the slow time dependence of the geodesic parameters $P^M(\tilde{w}) = \{E(\tilde{w}), L_z(\tilde{w}), Q(\tilde{w})\}$. The only remaining quasistatic effects are those associated with $\delta M(\tilde{w})$ and $\delta a(\tilde{w})$, which must be determined from the field equations. The strategy for determining these final remaining slowly varying will be presented in Section 5.3.

For the Lorenz gauge computation, the mass and spin parts of the metric perturbation do not appear in the relaxed Einstein field equation (5.17); the relaxed Einstein field operator $E_{\mu\nu}$ annihilates pure-mass or pure-spin contributions. Instead, the slow time evolution of the mass and spin is determined by the (now dynamical) fast-time average of the subleading Lorenz gauge condition,

$$\left\langle w_\mu \partial_{\tilde{w}} h^{(1)\mu\nu} \right\rangle = - \left\langle \Omega^{(1)A}(\tilde{w}) w_\mu \partial_A h^{(1)\mu\nu} \right\rangle . \tag{5.18}$$

Note that while the gauge condition (5.18) is computed at second order, it is independent of $h^{(2)\mathcal{P}}$, as the fast-time averaging opration removes all but the quasistatic part of the metric perturbation, and the leading covariant differentiation $\nabla^{(0)}$ then removes any dependence on $h^{(2)}$.

In the Teukolsky-Lousto-Campanelli formalism, we instead adhere more closely to the general procedure described in the multiscale text [350]. Generally, the slow evolution of the remaining parameters is determined by enforcing consistency with the multiscale construction: the nonlinear source which arises at second order must not give rise to solutions secular in the fast time variables q^A. The consistency constraint ensures that the nonlinear source, which involves slow time derivative contributions (see (5.14)), has no part which can be expressed as a homogeneous solution to the wave equation, and thereby fixes the slow time dependence of $\delta M(\tilde{w})$ and $\delta a(\tilde{w})$.

5.3 Teukolsky-Lousto-Campanelli wave equations in multiscale

5.3.1 The Teukolsky-Lousto-Campanelli wave equation

In this subsection, we reproduce the detailed form of the second-order scalar wave equations for the Teukolsky scalars. These equations were first derived in [356], but for completeness, we reproduce the resulting wave equations below in full. In addition, we compute the corrections to the Lousto-Campanelli source from slow time derivatives required for the second order multiscale wave equation. By necessity, our presentation of the Teukolsky-Lousto-Campanelli wave equations makes extensive use of the Newman-Penrose formalism. All required Newman-Penrose definitions are given in appendix 5.C.

Following the same steps as followed in [356], which in turn followed the steps of the original derivation of [355], we develop the pair of equations for ψ_4 and ψ_0. It is convenient to define several additional derivative operators,

$$d_1 = D - 4\rho - \bar{\rho} - 3\varepsilon + \bar{\varepsilon}, \tag{5.19a}$$

$$d_2 = \delta - 4\tau + \bar{\pi} - \bar{\alpha} - 3\beta, \tag{5.19b}$$

$$d_3 = \bar{\delta} + 3\alpha + \bar{\beta} + 4\pi - \bar{\tau}, \tag{5.19c}$$

$$d_4 = \Delta + 4\mu + \bar{\mu} + 3\gamma - \bar{\gamma}. \tag{5.19d}$$

Using these derivative operators, the general order-n wave equations for the Weyl scalars ψ_4 and ψ_0

are

$$
\begin{aligned}
&\rho^{-4}\Sigma(d_4^{(0)}(D+4\varepsilon-\rho)^{(0)}-d_3^{(0)}(\delta+4\beta-\tau)^{(0)}-3\psi_2^{(0)})(\psi_4^{(n)})\\
&=2\rho^{-4}\Sigma\cdot\sum_{p=1}^{n-1}\left[d_3^{(0)}(\delta+4\beta-\tau)^{(n-p)}-d_4^{(0)}(D+4\varepsilon-\rho)^{(n-p)}\right]\psi_4^{(p)}\\
&\quad-\left[d_3^{(0)}(\nabla+4\mu+2\gamma)^{(n-p)}-d_4^{(0)}(\bar{\delta}+4\pi+2\alpha)^{(n-p)}\right]\psi_3^{(p)}\\
&\quad-3\psi_2^{(0)}\left[(d_3-3\pi)^{(n-p)}\nu^{(p)}-(d_4-3\mu)^{(n-p)}\lambda^{(p)}\right]+3\left[d_3^{(0)}\nu^{(n-p)}-d_4^{(0)}\lambda^{(n-p)}\right]\psi_2^{(p)}\\
&\quad+4\pi d_4^{(0)}\Big[(\bar{\delta}-2\bar{\tau}+2\alpha)^{(n-p)}T_{n\bar{m}}^{(p)}-(\Delta+2\gamma-2\bar{\gamma}+\bar{\mu})^{(n-p)}T_{\bar{m}\bar{m}}^{(p)}\\
&\qquad\qquad-\lambda^{(n-p)}(T_{nl}+T_{m\bar{m}})^{(p)}+\bar{\sigma}^{(n-p)}T_{nn}^{(p)}+2\nu^{(n-p)}T_{l\bar{m}}^{(p)}\Big]\\
&\quad+4\pi d_3^{(0)}\Big[(\Delta+2\gamma+2\bar{\mu})^{(n-p)}T_{n\bar{m}}^{(p)}-(\bar{\delta}-\bar{\tau}+2\bar{\beta}+2\alpha)^{(n-p)}T_{nn}^{(p)}\\
&\qquad\qquad-\nu^{(n-p)}(T_{nl}+T_{m\bar{m}})^{(p)}-\bar{\nu}^{(n-p)}T_{\bar{m}\bar{m}}^{(p)}-\lambda^{(n-p)}T_{nm}^{(p)}\Big],
\end{aligned}
\tag{5.20}
$$

and

$$
\begin{aligned}
&(d_1^{(0)}(\Delta-4\gamma+\mu)^{(0)}-d_2^{(0)}(\bar{\delta}+\pi-4\alpha)^{(0)}-3\psi_2^{(0)})\psi_0^{(n)}\\
&=\sum_{p=1}^{n-1}\left[d_1^{(0)}(\Delta-4\gamma+\mu)^{(n-p)}-d_2^{(0)}(\bar{\delta}+\pi-4\alpha)^{(n-p)}\right]\psi_0^{(p)}\\
&\quad-\left[d_2^{(0)}(D-4\rho-2\varepsilon)^{(n-p)}-d_1^{(0)}(\bar{\delta}+\pi-4\alpha)^{(n-p)}\right]\psi_1^{(p)}\\
&\quad-3\psi_2^{(0)}\left[(d_1+3\rho)^{(n-p)}\sigma^{(p)}-(d_2+3\tau)^{(n-p)}\kappa^{(p)}\right]+3\left[d_1^{(0)}\sigma^{(n-p)}-d_2^{(0)}\kappa^{(n-p)}\right]\psi_2^{(p)}\\
&\quad+4\pi d_1^{(0)}\Big[(-D+2\varepsilon-2\bar{\varepsilon}+\bar{\rho})^{(n-p)}T_{mm}^{(p)}+(\delta+2\bar{\pi}-2\beta)^{(n-p)}T_{lm}^{(p)}\\
&\qquad\qquad\sigma^{(n-p)}(T_{lm}+T_{m\bar{m}})^{(p)}-2\kappa^{(n-p)}T_{nm}^{(p)}-\bar{\lambda}^{(n-p)}T_{ll}^{(p)}\Big]\\
&\quad+4\pi d_2^{(0)}\Big[(\delta+\bar{\pi}-2\bar{\alpha}-2\beta)^{(n-p)}T_{ll}^{(p)}+(-D+2\varepsilon+2\bar{\rho})^{(n-p)}T_{lm}^{(p)}\\
&\qquad\qquad+2\sigma^{(n-p)}T_{l\bar{m}}^{(p)}-\kappa^{(n-p)}(T_{ln}+T_{m\bar{m}})^{(p)}-\bar{\kappa}^{(n-p)}T_{mm}^{(p)}\Big].
\end{aligned}
\tag{5.21}
$$

Note that while the sources might appear to depend only on curvature scalars at various orders, the perturbed spin coefficients depend directly full metric perturbations at each lower order (given explicitly in Appendix 5.C). So, these equations practically allow the computation of the n^{th} order using scalars, provided computation or reconstruction of the metric perturbation for all orders $\leq (n-1)^{\text{th}}$.

As shown in the original derivation by [355], the differential operators on the left-hand side of the pair of separable equations (5.21) and (5.20) can be re-written using the Kinnersly tetrad and in Boyer-Lindquist coordinates,

$$
\begin{aligned}
&(d_1^{(0)}(\Delta-4\gamma+\mu)^{(0)}-d_2^{(0)}(\bar{\delta}+\pi-4\alpha)^{(0)}-3\psi_2^{(0)})\psi_0^{(n)}\\
&=-\Bigg\{\left[\frac{\varpi^4}{\Delta}-a^2\sin^2(\theta)\right]\partial_t^2+\frac{4Mar}{\Delta}\partial_t\partial_\phi+\left[\frac{a^2}{\Delta}-\frac{1}{\sin^2(\theta)}\right]\partial_\phi^2\\
&\quad-\Delta^{-2}\partial_r(\Delta^3\partial_r)-\frac{1}{\sin(\theta)}\partial_\theta(\sin(\theta)\partial_\theta)-4\left[\frac{a(r-M)}{\Delta}+\frac{i\cos(\theta)}{\sin^2(\theta)}\right]\partial_\phi\\
&\quad-4\left[\frac{M\varpi^2}{\Delta}-r-ia\cos(\theta)\right]\partial_t+(4\cot(\theta)-2)\Bigg\}\psi_0^{(n)},
\end{aligned}
\tag{5.22a}
$$

$$
\begin{aligned}
&\rho^{-4}\Sigma(d_4^{(0)}(D+4\varepsilon-\rho)^{(0)}-d_3^{(0)}(\delta+4\beta-\tau)^{(0)}0-3\psi_2^{(0)})(\psi_4^{(n)})\\
&=-\Bigg\{\left[\frac{\varpi^4}{\Delta}-a^2\sin^2(\theta)\right]\partial_t^2+\frac{4Mar}{\Delta}\partial_t\partial_\phi+\left[\frac{a^2}{\Delta}-\frac{1}{\sin^2(\theta)}\right]\partial_\phi^2\\
&\quad-\Delta^{2}\partial_r(\Delta^{-1}\partial_r)-\frac{1}{\sin(\theta)}\partial_\theta(\sin(\theta)\partial_\theta)+4\left[\frac{a(r-M)}{\Delta}+\frac{i\cos(\theta)}{\sin^2(\theta)}\right]\partial_\phi\\
&\quad+4\left[\frac{M\varpi^2}{\Delta}-r-ia\cos(\theta)\right]\partial_t+(4\cot(\theta)+2)\Bigg\}\left(\rho^{-4}\psi_0^{(n)}\right).
\end{aligned}
\tag{5.22b}
$$

For the multiscale expansion, we wish to write the homogeneous wave operators (5.22) in terms of the modified Boyer-Lindquist coordinates in which we use the time coordinate $w=t+h(r)$, and the background takes takes the coordinate form (5.5). Using the coordinate transformation, we derive the Kinnersly tetrad in the modified Boyer-Lindquist coordinates,

$$
\vec{l}=\left(\frac{\varpi^2}{\Delta}+h'(r)\right)\partial_w+\partial_r+\frac{a}{\Delta}\partial_\phi,
\tag{5.23a}
$$

$$
\vec{n}=\left(\frac{\varpi^2}{2\Sigma}-\frac{\Delta}{2\Sigma}h'(r)\right)\partial_w-\frac{\Delta}{2\Sigma}\partial_r+\frac{a}{2\Sigma}\partial_\phi,
\tag{5.23b}
$$

$$
\vec{m}=\frac{1}{\sqrt{2}(r+ia\cos(\theta))}\left(ia\sin(\theta)\partial_w+\partial_\theta+\frac{i}{\sin(\theta)}\partial_\phi\right),
\tag{5.23c}
$$

$$
\vec{\bar{m}}=\frac{1}{\sqrt{2}(r-ia\cos(\theta))}\left(-ia\sin(\theta)\partial_w+\partial_\theta-\frac{i}{\sin(\theta)}\partial_\phi\right),
\tag{5.23d}
$$

and the resulting coordinate-transformed differential operators acting on the Teukolsky variables

are

$$
\begin{aligned}
&(d_1^{(0)}(\Delta - 4\gamma + \mu)^{(0)} - d_2^{(0)}(\bar{\delta} + \pi - 4\alpha)^{(0)} - 3\psi_2^{(0)})\psi_0^{(n)} \\
&= \Bigg\{ \left[\frac{\varpi^4}{\Delta} - a^2\sin^2(\theta) - \Delta h'(r)^2\right]\partial_w^2 + \frac{4Mar}{\Delta}\partial_w\partial_\phi + \left[\frac{a^2}{\Delta} - \frac{1}{\sin^2(\theta)}\right]\partial_\phi^2 - \Delta^{-2}\partial_r(\Delta^3\partial_r) \\
&\quad - \Delta^{-2}\partial_r(\Delta^3 h'(r)\partial_w) - \Delta h'(r)\partial_w\partial_r - \frac{1}{\sin(\theta)}\partial_\theta(\sin(\theta)\partial_\theta) - 4\left[\frac{a(r-M)}{\Delta} + \frac{i\cos(\theta)}{\sin^2(\theta)}\right]\partial_\phi \\
&\quad - 4\left[\frac{M\varpi^2}{\Delta} - r - ia\cos(\theta)\right]\partial_w + (4\cot(\theta) - 2)\Bigg\}\psi_0^{(n)},
\end{aligned}
\tag{5.24a}
$$

and

$$
\begin{aligned}
&\rho^{-4}\Sigma(d_4^{(0)}(D + 4\varepsilon - \rho)^{(0)} - d_3^{(0)}(\delta + 4\beta - \tau)^{(0)}0 - 3\psi_2^{(0)})(\psi_4^{(n)}) \\
&= \Bigg\{ \left[\frac{\varpi^4}{\Delta} - a^2\sin^2(\theta) - \Delta h'(r)^2\right]\partial_w^2 + \frac{4Mar}{\Delta}\partial_w\partial_\phi + \left[\frac{a^2}{\Delta} - \frac{1}{\sin^2(\theta)}\right]\partial_\phi^2 - \Delta^2\partial_r(\Delta^{-1}\partial_r) \\
&\quad - \Delta^2\partial_r(\Delta^{-1}h'(r)\partial_w) - \Delta h'(r)\partial_w\partial_r - \frac{1}{\sin(\theta)}\partial_\theta(\sin(\theta)\partial_\theta) + 4\left[\frac{a(r-M)}{\Delta} + \frac{i\cos(\theta)}{\sin^2(\theta)}\right]\partial_\phi \\
&\quad + 4\left[\frac{M\varpi^2}{\Delta} - r - ia\cos(\theta)\right]\partial_w + (4\cot(\theta) + 2)\Bigg\}\left(\rho^{-4}\psi_4^{(n)}\right).
\end{aligned}
\tag{5.24b}
$$

The remaining quantities involved in the wave equations (5.20) and (5.21) are the first-order perturbed tetrads, spin coefficients, and curvature scalars. The full set of first order quantities in terms of the first order metric perturbations are given in Appendix 5.C. A second order Teukolsky-Lousto-Campanelli computation will have the iterative procedure of evaluating the first order Teukolsky equation, then reconstructing the first order metric perturbation as described in 5.3.3, and finally using those first order metric perturbations in the source terms of the second order TLC wave equations (5.20) and (5.21).

The promotion of the TLC wave equations (5.20), (5.21) to a multiscale version requires only the expansion of the multiscale time derivatives,

$$
\partial_w \to \Omega^A\partial_{q^A} + \varepsilon\partial_{\tilde{w}} + \varepsilon\Omega^{A(1)}\partial_{q^A} + \mathcal{O}(\varepsilon^2) \tag{5.25}
$$

The remaining details of the original wave equations carry through. In addition to the original second order sources, though, we require additional multiscale sources. We denote by ${}_{-2}S_{TT}^{(2)}[\psi_4^{(1)}]$ the multiscale source for the second-order $\psi_4^{(2)}$, which depends only on $\psi_4^{(1)}$. Correspondingly, there is a ${}_{+2}S_{TT}^{(2)}[\psi_0^{(1)}]$, which is the analogous source for the ψ_0 spin $+2$ equation. Each of these sources

is added directly to the right-hand side of the wave equations (5.20) and (5.21). These sources are extracted from the adjusted Teukolsky wave operators (5.24) by expanding the multiscale derivatives according to (5.25),

$$
\begin{aligned}
{}_{+2}S^{(2)}_{TT}[\psi^{(1)}_0] =& \Bigg\{ \left[\frac{1}{\Delta} - a^2\sin^2(\theta) - \Delta h'(r)^2\right]\left(2\Omega^A\partial_{q^A}\partial_{\tilde{w}} + 2\Omega^A\Omega^{B(1)}\partial_{q^A}\partial_{q^B} + \Omega'^A(\tilde{w})\partial_{q^A}\right) \\
&+ \left[\frac{4Mar}{\Delta}\partial_\phi - \Delta^{-2}\partial_r(\Delta^3 h'(r)) - 2\Delta h'\partial_r \right. \\
&\qquad \left. - 4\left(\frac{M\varpi^2}{\Delta} - r - ia\cos(\theta)\right)\right]\left(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_{q^A}\right)\Bigg\}\psi^{(1)}_0 \qquad (5.26a) \\
{}_{-2}S^{(2)}_{TT}[\psi^{(1)}_4] =& \Bigg\{ \left[\frac{1}{\Delta} - a^2\sin^2(\theta) - \Delta h'(r)^2\right]\left(2\Omega^A\partial_{q^A}\partial_{\tilde{w}} + 2\Omega^A\Omega^{B(1)}\partial_{q^A}\partial_{q^B} + \Omega'^A(\tilde{w})\partial_{q^A}\right) \\
&+ \left[\frac{4Mar}{\Delta}\partial_\phi - \Delta^{2}\partial_r(\Delta^{-1} h'(r)) - 2\Delta h'\partial_r \right. \\
&\qquad \left. + 4\left(\frac{M\varpi^2}{\Delta} - r - ia\cos(\theta)\right)\right]\left(\partial_{\tilde{w}} + \Omega^{(1)A}\partial_{q^A}\right)\Bigg\}\left(\rho^{-4}\psi^{(1)}_4\right) \qquad (5.26b)
\end{aligned}
$$

The Weyl scalars $\psi^{(2)}_0$ and $\psi^{(2)}_4$ are neither gauge nor tetrad invariant. Under first order gauge transformations, both quantities receive corrections dependent on the zeroth and first order Weyl scalars. However, as pointed out by [356], the asymptotic waveform remains well-defined as long as it is computed in an appropriately asymptotically flat gauge. Further, for computation of the post-adiabatic waveform, from the second-order solution we require only the gauge-invariant slow evolution of the orbital parameters at subleading order to determine the subleading frequency corrections.

For the purposes of explicit comparison between self force computations and numerical relativity or post-Newtonian computations, it is desirable to have a method for constructing gauge-invariant quantities which parameterize the dynamics of the system through second order. A method for accomplishing this was also suggested in [356], although the details of the computation presented in that work assumed the existence of a complicated tertiary gauge. In general, if a gauge can be completely specified, one can construct an explicit gauge vector $\xi^\mu[h^{(1)}]$ such that, in any gauge, the quantity

$$h'_{\mu\nu} = h_{\mu\nu} + 2\nabla_{(\mu}\xi_{\nu)} \qquad (5.27)$$

takes the value of $h_{\mu\nu}$ in a particular, chosen gauge. The gauge vector $\xi^\mu[h^{(1)}]$ may then be used to construct gauge-invariant Weyl scalars $\psi^{(2)}_0$ and $\psi^{(2)}_4$. It is worth emphasizing that such a quantity

is genuinely gauge invariant, despite the fact that the values will certainly depend on which gauge is chosen to use in the construction of $\xi^\mu[h^{(1)}]$. The correct interpretation is that there is a family of gauge-invariant quantities, each of which takes the value associated with a particular gauge choice in every gauge.

5.3.2 Modes of the Teukolsky-Lousto-Campanelli variables

While the nonlinear wave equations which govern the second order perturbation have a source which is significantly more complicated than the first order wave source, homogeneous modes of the wave equation themselves are no more intricate. The wave operator which acts on the second order field, as discussed in Section 5.3.1, is identical to that which operates on the first order field. In this section, we will review the Teukolsky modes for generic sources with the understanding that it applies to both the first and second order perturbations.

The homogeneous Teukolsky wave equation (5.24) is a separable wave equation, and is best expressed as modes in terms of the spin-weighted moments ($s = 2$ for ψ_0, $s = -2$ for ψ_4), frequencies ω, and spheroidal harmonic constants l and m,

$$\Psi_s = {}_sR_{\omega lm}(r)\,{}_s\Theta_{\omega lm}(\theta)e^{im\phi}e^{-i\omega w}, \tag{5.28}$$

where the spin weighted spheroidal harmonic functions ${}_s\Theta_{\omega lm}(\theta)$ are defined to satisfy

$$\begin{aligned}
&{}_s\lambda_{\omega lm}\,{}_s\Theta_{\omega lm}(\theta)\\
&= \frac{1}{\sin(\theta)}\frac{d}{d\theta}\left(\sin\theta\frac{d_s\Theta_{\omega lm}(\theta)}{d\theta}\right)\\
&\quad + \left(a^2\omega^2\cos^2(\theta) - \frac{m^2}{\sin^2(\theta)} - 2a\omega s\cos(\theta) - \frac{2ms\cos(\theta)}{\sin^2(\theta)} - s^2\cot^2(\theta) - s^2\right)\Theta_{\omega lm}(\theta).
\end{aligned} \tag{5.29}$$

Then, the radial Teukolsky modes must obey the remaining components of the wave equation:

$$\begin{aligned}
&-\Delta^{-s}\partial_r(\Delta^{s+1}\partial_r{}_sR_{\omega lm}(r)) + i\Delta^{-s}\omega\partial_r(\Delta^{s+1}{}_sR_{\omega lm}(r)) + i\Delta\omega\partial_r{}_sR_{\omega lm}(r)\\
&+ {}_s\lambda_{\omega lm}\,{}_sR_{\omega lm}(r) - \omega^2\left[\frac{\varpi^4}{\Delta} - a^2\sin^2(\theta) - \Delta h'(r)\right]{}_sR_{\omega lm}(r)\\
&- 4i\omega\left[\frac{M\varpi^2}{\Delta} - r - ia\cos(\theta)\right] - 4im\left(\frac{a(r-M)}{\Delta}\right){}_sR_{\omega lm}(r) = -{}_sT_{\omega lm}(r),
\end{aligned} \tag{5.30}$$

where

$${}_sT_{\omega lm}(r) = \int dwd\Omega e^{i\omega w}e^{-im\phi}\,{}_sS^*_{\omega lm}(\theta,\phi)\,{}_sT(w,r,\theta,\phi), \tag{5.31}$$

and [355, 357],

$$
\begin{aligned}
&{}_2T(w,r,\theta,\phi)\\
&\quad= 2\Big\{(\delta+\bar\pi-\bar\alpha-3\beta-4\tau)\left[(D-2\varepsilon-2\bar\rho)T_{lm}-(\delta+\bar\pi-2\bar\alpha-2\beta)T_{ll}\right]\\
&\qquad+(D-3\varepsilon+\bar\varepsilon-4\rho-\bar\rho)\left[(\delta+2\bar\rho-2\beta)T_{lm}-(D-2\varepsilon+2\bar\varepsilon-\bar\rho)T_{mm}\right]\Big\},
\end{aligned}
\tag{5.32a}
$$

$$
\begin{aligned}
&{}_{-2}T(w,r,\theta,\phi)\\
&\quad= 2\rho^{-4}\Big\{(\Delta+3\gamma-\bar\gamma+4\mu+\bar\mu)\left[(\bar\delta-2\bar\tau+2\alpha)T_{m\bar m}-(\Delta+2\gamma-2\bar\gamma+\bar\mu)T_{\bar m\bar m}\right]\\
&\qquad+(\bar\delta-\bar\tau+\bar\beta+3\alpha+4\pi)\left[(\Delta+2\gamma+2\bar\mu)T_{n\bar m}-(\bar\delta-\bar\tau+2\bar\beta+2\alpha)T_{nn}\right]\Big\}.
\end{aligned}
\tag{5.32b}
$$

The modes which satisfy (5.30) can be classified by their asymptotic dependence. The solutions have sufficient freedom such that they can be chosen to vanish at any one of the past horizon H^-, the future horizon H^+, past null infinity $\mathcal{I}^-$, or future null infinity $\mathcal{I}^+$. These choices of incoming or outgoing boundary conditions at each of the null horizons of the spacetime define the natural mode decompositions of the radial equation (5.30). The "in" modes vanish on H^-, the "up" modes vanish on $\mathcal{I}^-$, the "out" modes vanish on H^+, and the "down" modes vanish on $\mathcal{I}^+$. The pairs of modes "in" and "up", or "out" and "down" each form a basis for the the complete set of solutions to the Teukolsky wave equation.

The full set of mode functions is denoted,

$$
{}_s\Psi^{\text{in}}_{\omega lm}(w,r,\theta,\phi) = e^{-i\omega w}\,{}_sR^{\text{in}}_{\omega lm}(r)\,{}_sS_{\omega lm}(\theta,\phi) \tag{5.33a}
$$

$$
{}_s\Psi^{\text{out}}_{\omega lm}(w,r,\theta,\phi) = e^{-i\omega w}\,{}_sR^{\text{out}}_{\omega lm}(r)\,{}_sS_{\omega lm}(\theta,\phi) \tag{5.33b}
$$

$$
{}_s\Psi^{\text{up}}_{\omega lm}(w,r,\theta,\phi) = e^{-i\omega w}\,{}_sR^{\text{up}}_{\omega lm}(r)\,{}_sS_{\omega lm}(\theta,\phi) \tag{5.33c}
$$

$$
{}_s\Psi^{\text{down}}_{\omega lm}(w,r,\theta,\phi) = e^{-i\omega w}\,{}_sR^{\text{down}}_{\omega lm}(r)\,{}_sS_{\omega lm}(\theta,\phi). \tag{5.33d}
$$

The mode functions then require slight alteration to obey the multiscale promoted wave equations. The multiscale modes are now entries in a Fourier series over the 2π-periodic angle variables

q^A, rather than the Fourier transform used for arbitrary time dependence,

$$
\begin{aligned}
{}_s\Psi^{\text{in}}_{\omega lm}(q^A, r, \theta, \phi) =& e^{-ik_A q^A}\, {}_sR^{\text{in}}_{\omega lm}(r)\, {}_sS_{\omega lm}(\theta, \phi) && \text{(5.34a)}\\
{}_s\Psi^{\text{out}}_{\omega lm}(q^A, r, \theta, \phi) =& e^{-ik_A q^A}\, {}_sR^{\text{out}}_{\omega lm}(r)\, {}_sS_{\omega lm}(\theta, \phi) && \text{(5.34b)}\\
{}_s\Psi^{\text{up}}_{\omega lm}(q^A, r, \theta, \phi) =& e^{-ik_A q^A}\, {}_sR^{\text{up}}_{\omega lm}(r)\, {}_sS_{\omega lm}(\theta, \phi) && \text{(5.34c)}\\
{}_s\Psi^{\text{down}}_{\omega lm}(q^A, r, \theta, \phi) =& e^{-ik_A q^A}\, {}_sR^{\text{down}}_{\omega lm}(r)\, {}_sS_{\omega lm}(\theta, \phi). && \text{(5.34d)}
\end{aligned}
$$

Furthermore, the mode expansion must also include slow time $\tilde{w}$ dependent mode amplitudes,

$$
\begin{aligned}
\psi_0 &= \sum_{k^A lm} \left[{}_2X^{\text{out}}_{k^A lm}(\tilde{w})\, {}_2R^{\text{out}}_{\omega lm}(r) + {}_2X^{\text{down}}_{k^A lm}(\tilde{w})\, {}_2R^{\text{down}}_{\omega lm}(r) \right] e^{-ik_A q^A}\, {}_2S_{\omega lm}(\theta, \phi) && \text{(5.35a)}\\
&= \sum_{k^A lm} \left[{}_2X^{\text{in}}_{k^A lm}(\tilde{w})\, {}_2R^{\text{in}}_{\omega lm}(r) + {}_2X^{\text{up}}_{k^A lm}(\tilde{w})\, {}_2R^{\text{up}}_{\omega lm}(r) \right] e^{-ik_A q^A}\, {}_2S_{\omega lm}(\theta, \phi) && \text{(5.35b)}\\
\rho^{-4}\psi_4 &= \sum_{k^A lm} \left[{}_{-2}X^{\text{out}}_{k^A lm}(\tilde{w})\, {}_{-2}R^{\text{out}}_{\omega lm}(r) + {}_{-2}X^{\text{down}}_{k^A lm}(\tilde{w})\, {}_{-2}R^{\text{down}}_{\omega lm}(r) \right] e^{-ik_A q^A}\, {}_{-2}S_{\omega lm}(\theta, \phi) && \text{(5.35c)}\\
&= \sum_{k^A lm} \left[{}_{-2}X^{\text{in}}_{k^A lm}(\tilde{w})\, {}_{-2}R^{\text{in}}_{\omega lm}(r) + {}_{-2}X^{\text{up}}_{k^A lm}(\tilde{w})\, {}_{-2}R^{\text{up}}_{\omega lm}(r) \right] e^{-ik_A q^A}\, {}_{-2}S_{\omega lm}(\theta, \phi), && \text{(5.35d)}
\end{aligned}
$$

where we denote the slow-time dependent mode coefficients with ${}_sX^{<\text{mode}>}_{k^A lm}(\tilde{w})$.

5.3.3 Reconstruction of first order modes

The flux formalism introduced in Section 5.4 offers the ability to compute the effects of the dissipative second order self force from the second-order Teukolsky variables, which is all that will be required of the second order self force for post-adiabatic waveforms. However, the form of the Teukolsky-Lousto-Campanelli equations (5.20,5.21) offers little hope for any method which avoids computing the full first order metric, as the first order Newman-Penrose spin coefficients depend directly on all components of the first order metric perturbation (see Appendix 5.C).

It will be necessary to reconstruct the radiative modes via the CCK procedure [357, 358]. We review here the details we require for the metric reconstruction formalism. In addition to the reconstructed radiative modes, we will require the $l = 0$ and $l = 1$ completion parts of the first order metric perturbations which are not derivable from the Chrzanowski procedure [359, 360]. These procedures have been successfully implemented for full first order self force computations in Kerr [313, 361]. Additionally, however, beyond the reconstruction and completion there will be

unfixed functions of slow time $\tilde{w}$ which parameterize the slow variation of the background mass and spin $\delta M(\tilde{w}), \delta a(\tilde{w})$, which will be obtained by imposing consistency of the second-order source with the multiscale construction, discussed in detail in Section 5.3.5.

There is a mild modification of the CCK procedure required for the multiscale formalism, associated with reconstructing instead of the physical metric perturbation $h(w, r, \theta, \phi)$ from $\psi_{0/4}(w, r, \theta, \phi)$, the multiscale metric perturbation $h(q^A, \tilde{w}, r, \theta, \phi)$ from $\psi_{0/4}(q^A, \tilde{w}, \theta, \phi)$. The wave equations which the metric perturbation satisfies are those associated with the wave operator dependent only on the fast time variable. Therefore, the reconstruction procedure proceeds directly, except with the replacement of all time derivatives ∂_w with $\Omega^{(0)A}\partial_A$, which excludes any slow-time derivatives. Ultimately, the distinction is immaterial at first order, as any corrections associated with the slow time variables would be contributions to the reconstruction at second order, which is not necessary and we do not attempt.

The metric perturbations are reconstructed by first deriving the Hertz potential Ψ, which satisfies the same Teukolsky equation as the spin -2 part of the Teukolsky wave equations, which are schematically written as

$$W_{\pm 2}[\psi_{\pm 2}] = 4\pi\Sigma T_{\pm 2}, \tag{5.36}$$

where $\psi_{+2} = \psi_0$ and $\psi_{-2} = \rho^{-4}\psi_4$. The Hertz potential also satisfies the differential equations [362–364],

$$\psi_0 = \frac{1}{2}(D - 3\varepsilon + \bar{\varepsilon} - \bar{\rho})(D - 2\varepsilon + 2\bar{\varepsilon} - \bar{\rho})(D - \varepsilon + 3\bar{\varepsilon} - \bar{\rho})(D + 4\bar{\varepsilon} + 3\bar{\rho})\bar{\Psi}, \tag{5.37a}$$

$$\begin{aligned}\psi_4 = &\frac{1}{2}(\bar{\delta} + 3\alpha + \bar{\beta} - \bar{\tau})(\bar{\delta} + 2\alpha + 2\bar{\beta} - \bar{\tau})(\bar{\delta} + \alpha + 3\bar{\beta} - \bar{\tau})(\bar{\delta} + 4\bar{\beta} + 3\bar{\tau})\bar{\Psi} \\ &+ \frac{3}{2}\psi_2^{(0)}\left[\tau(\bar{\delta} + 4\alpha) - \rho(\Delta + 4\gamma) - \mu(D + 4\varepsilon) + \pi(\delta + 4\beta) + 2\psi_2^{(0)}\right]\Psi,\end{aligned} \tag{5.37b}$$

where all time derivatives in (5.37) are expressed as the leading multiscale analogs $\partial_t \to \Omega^{(0)A}\partial_A$. The remaining details of the solutions to these fourth-order differential equations in the Kerr vacuum are given in [365], and generalize directly to strictly periodic functions of the time variables q^A. From this Hertz potential, the metric perturbation may be reconstructed by applying the differential operator introduced by Chrzanowski [357, 358], which we will follow [365] in notating schematically,

$$h_{\alpha\beta}^{\mathrm{IRG/ORG}} = \Pi_{\alpha\beta}^{\mathrm{IRG/ORG}}[\Psi]. \tag{5.38}$$

The reconstruction operator $\Pi^{\rm IRG}$ takes the explicit form [364],

$$\begin{aligned}\Pi^{\rm IRG}_{\alpha\beta}[\Psi] = -\Big\{ & l_\alpha l_\beta \left[(\bar\delta + \alpha + 3\bar\beta - \bar\tau)(\bar\delta + 4\bar\beta + 3\bar\tau) - \lambda(D + 4\bar\varepsilon + 3\bar\rho)\right] \\ & \bar m_\alpha \bar m_\beta (D - \varepsilon + 3\bar\varepsilon - \bar\rho)(D + 4\bar\varepsilon + 3\bar\rho) \\ & - l_{(\alpha}\bar m_{\beta)}\big[(D + \varepsilon + 3\bar\varepsilon + \rho - \bar\rho)(\bar\delta + 4\beta + 3\bar\tau) \\ & + (\bar\delta - \alpha + 3\bar\beta - \pi - \bar\tau)(D + 4\bar\varepsilon + 3\bar\rho)\big]\Big\}\bar\Psi + (\text{complex conjugate}),\end{aligned} \tag{5.39}$$

for ingoing radiation gauge, which transparently satisfies the gauge condition $h_{\alpha\beta}l^\alpha = 0$. The analogous operator for reconstructing the first order metric perturbation in outgoing radiation gauge is obtained by exchanging the l^α and n^α null vectors.

The metric reconstruction procedure can be used to find all $l \geq 2$ modes of the metric perturbation, at all radii which do not share a spherical shell with regions of nonvanishing stress energy $T_{\alpha\beta}$. For the first-order field, the result is a discontinuous feature in the radial functions in the mode decomposition at the instantaneous radius of the small body. For general orbits, the radius at which the discontinuity occurs evolves with time on the orbital timescale, and presents a significant computational problem for mode decomposition. The mode decomposition of any non-smooth function involves comparatively large amplitudes of high frequency modes, causing spectral approximation methods to converge slowly. This problem with discontinuities is known as Gibbs phenomenon [366], and past work has explored methods of side-stepping the problem for EMRIs. To present a complete method of computing inspirals, we will advocate the use of the method of extended homogeneous solutions [367], for which the application to second order multiscale approximations is discussed in Section 5.3.4.

The metric reconstruction procedure fails to reproduce the $l = 0$ and $l = 1$ parts of the metric perturbation not encoded in the ψ_0 and ψ_4 fields. These degrees of freedom are referred to as the "completion" part of the metric perturbation, and their derivation is detailed in [313,360,368]. The full metric perturbation inside (-) and outside (+) the orbit is given by

$$h^\pm_{\alpha\beta} = h^{{\rm rec}\pm}_{\alpha\beta} + \mathcal{E}^\pm h^{(\delta M)}_{\alpha\beta} + \mathcal{J}^\pm h^{(\delta J)}_{\alpha\beta}, \tag{5.40}$$

where $\mathcal{E}^+ - \mathcal{E}^- = E$ and $\mathcal{J}^+ - \mathcal{J}^- = L_z$, and E and L_z are the orbital parameters. The additional required metric perturbations can be obtained in a background-compatible gauge by varying the

background metric with respect to M and J.

$$h^{(\delta M)}_{\alpha\beta} = \frac{\partial g_{\alpha\beta}(x^\mu; M, J)}{\partial M}, \tag{5.41a}$$

$$h^{(\delta J)}_{\alpha\beta} = \frac{\partial g_{\alpha\beta}(x^\mu; M, J)}{\partial J}, \tag{5.41b}$$

where $J = aM$. The final remaining contribution to be determined is then an overall mass and spin contribution, which is permitted to depend on slow-time variable $\tilde{w}$. In the self force treatment which does not use the multiscale expansion [360], the remaining degree of freedom can be found via asymptotic constraints of total mass and angular momentum. For the multiscale expansion, in which the mass and angular momentum are dynamical and must are fixed by subleading slow-time dependent equations of motion, we will require an alternative strategy. The methods we use for finding the mass and angular momentum for the TLC wave equation strategy are discussed in Section 5.3.5.

5.3.4 Multiscale TLC computations using extended solutions

We turn now to a discussion of the practical problem of the convergence of a frequency and spherical harmonic spectral decomposition for the Teukolsky-Lousto-Campanelli equations for EMRIs. The trouble for convergence of spectral decompositions is the discontinuity caused by the pointlike source of the small companion, which gives rise to sharp behavior in the radial dependence of all modes of the spherical harmonic decomposition. The spectral decomposition then converges at sub-exponential rate, making computation of the inspiral expensive. The solution to this problem is to separately compute multiple solutions which are smooth across the discontinuous source, and assemble a physical solution by transitioning abruptly between those solutions. For homogeneous wave equations, the technique is known as the method of extended homogeneous solutions [367,369], and the extension to particular equations is known as the method of extended particular solutions [370].

For each mode in the decomposition, the method of homogeneous solutions assigns a separate amplitude ${}_sZ^{\pm}_{\omega lm}$ to the solutions inside the orbit ($-$) and to solutions outside the orbit ($+$). The physical solution is then taken to be the sum of modes weighted by mode amplitudes ${}_sZ^{+}_{\omega lm}$ for all points in spacetime instantaneously outside the orbit, and the sum of modes weighted by amplitudes

${}_sZ^{-}_{\omega lm}$ for all points instantaneously within the orbit. It is notable that, in this context, the source only appears in determining the transition between the inside and outside solutions.

The method of extended particular solutions functions in much the same way, but is adapted to derive the radially varying mode weights ${}_sZ^{\pm}_{\omega lm}(r)$ via the method of variation of parameters. The extension to particular solutions assumes a smooth source at all points away from a single shell at which the solution is expected to have a sharp feature. The mode amplitudes are then obtained by a radial integral via variation of parameters, integrated either from $r^* \to +\infty$ for the $(+)$ modes, or from $r^* \to -\infty$ for the $(-)$ modes. Finally, similar to the method of extended homogeneous solutions, the physical solution is obtained after re-summing the mode decompositions and transitioning between the inner and outer solutions at the instantaneous radius of of the small companion.

These methods may be readily adopted for use in the first and second order expansions in the multiscale context. We propose the following computational workflow to avoid Gibbs phenomena in the multiscale evaluation of TLC wave equations:

- Evaluate the retarded Teukolsky modes for the physical pointlike source using the method of extended homogeneous solutions.
- Reconstruct the metric perturbation in the Ingoing and Outgoing radiation gauge, as appropriate in the regions inside and outside the orbit.
- Evaluate the residual metric perturbation at all relevant points using analytic form of the puncture metric.
- Evaluate the resulting second-order source at points inside and outside the instantaneous orbit for use in the method of particular solutions.
- Evaluate the inner and outer second order TLC wave solution using the method of extended particular solutions.

The final result will give a pair of second order TLC modes separately valid instantaneously inside the orbit and outside the orbit of the small companion. Those modes can then be used to conclude

information about dynamical invariants in the interaction zone, or propagated out to null horizons to infer flux values as discussed in Section 5.4.

5.3.5 Slow time evolution of mass and spin in radiation gauge

In this section, we give the generalized procedure for finding the slow evolution of the mass and the spin of the spacetime as gravitational radiation is accreted. For this derivation, we do not assume the Lorenz gauge, and will instead take advantage of the $l = 0$ part of the Ricci tensor to constrain the slowly evolving parameters of the spacetime. In the first subsection, we will demonstrate the general procedure by considering a pure mass perturbation in Schwarzschild spacetime. Building on the construction, we then derive in Section 5.3.5.2 the slow evolution of both the mass and the spin in a Kerr background.

5.3.5.1 Mass evolution in Schwarzschild

As a toy example to demonstrate our procedure, consider the slow evolution of the Schwarzschild mass, and assume that the net spin accretion is sufficiently small to neglect the corresponding δa. The ansatz for the perturbed metric is then,

$$h^{(1)}_{\alpha\beta} = \frac{\partial g^{(0)}_{\alpha\beta}}{\partial M}\delta M^{(1)}(\tilde{w}) + \mathcal{F}^{(1)}_{\alpha\beta}[\bar{x}^i, q^A, P^{(0)M}(\tilde{w})], \tag{5.42a}$$

$$h^{(2)}_{\alpha\beta} = \frac{\partial g^{(0)}_{\alpha\beta}}{\partial M}\delta M^{(2)}(\tilde{w}) + \mathcal{F}^{(2)}_{\alpha\beta}[x^i, q^A, P^{(0)M}(\tilde{w})], \tag{5.42b}$$

We additionally take the gauge for which $\mathcal{F}_{\alpha\beta}$ is in the convenient gauge for metric reconstruction, either ingoing or outgoing radiation gauge, and the remaining $l = 0$ part associated with mass evolution is in Schwarzschild gauge (such that (5.42) holds for the Schwarzschild coordinate metric $g^{(0)}_{\alpha\beta}$). This type of gauge choice is often used when discussing completion parts of the metric reconstruction [360]. Finally, we also impose that the quasistatic $l = 0$ part of $\mathcal{F}_{\alpha\beta}$ vanishes for

radii within the innermost point of the orbit, so that

$$\int d^3q^A \int d^2\Omega R^{(1)}_{\alpha\beta}[\mathcal{F}^{(1)}]\Big|_{r<r^-} = 0$$
$$\int d^3q^A \int d^2\Omega R_{\alpha\beta}[\mathcal{F}^{(2)}]\Big|_{r<r^-} = 0$$
$$\int d^3q^A \int d^2\Omega R_{\alpha\beta}[\mathcal{F}^{(1)}]\Big|_{r<r^-} = 0. \tag{5.43a}$$

We emphasize, however, that $\mathcal{F}$ will contain nontrivial $l = 0$ contributions outside the periapse of the small companion.

Proceeding with the computation under this gauge, the perturbed metric $\partial_M g^{(0)}$ has only three nonvanishing components:

$$\delta_M g^{(0)}_{ww} = \frac{2}{r}, \tag{5.44a}$$
$$\delta_M g^{(0)}_{rw} = -\frac{2h'(r)}{r}, \tag{5.44b}$$
$$\delta_M g^{(0)}_{rr} = \frac{2r}{(r-2M)^2} + \left(h'(r)\right)^2 \frac{2}{r}. \tag{5.44c}$$

The only nonvanishing component of the spherically averaged Ricci tensor contribution $R^{(1)l=0}_{\alpha\beta}[h^{(1)}]$ is then,

$$\int d^3q^A \int d^2\Omega R^{(1)}_{rt}[h^{(1)}]\Big|_{r<r^-} = \frac{2\partial_{\tilde{t}}\delta M}{-2Mr + r^2}. \tag{5.45}$$

The Einstein field equation at second order can be written as 5.12:

$$R^{(1)}_{ab}[h^{(2)}] = -R^{(1)}_{ab}[h^{(1)}] - R_{ab}[h^{(1)}, h^{(1)}] \tag{5.46}$$

The left-hand side of this equation will depend linearly on the second-order perturbed metric. We note from the dependence (5.45) that the slow-time evolution of the Schwarzschild mass at leading order is determined entirely by the rt component of the Einstein field equation. The contribution of the second order mass variation to the second order Einstein field equation vanishes, as (5.46) will depend only on the fast time derivative of any such mass variation term.

Taking an average over the fast time variables of the $l = 0$ part of the Einstein field equation (5.46), at a sphere of radius $r < r^-$, we find:

$$\begin{aligned}\partial_{\tilde{t}}\delta M =& -\frac{r^2 - 2Mr}{2(2\pi)^3(4\pi)} \int d^3q^A \int d^2\Omega\, R_{tr}[\mathcal{F}, \mathcal{F}],\\ =& \frac{-r^2}{4\pi(2\pi)^3} \int d^3q^A \int dS_{\alpha\beta}\xi^\alpha_{(t)} \left(R^\beta{}_\gamma[\mathcal{F}, \mathcal{F}]\right) \xi^\gamma_{(t)}.\end{aligned} \tag{5.47}$$

This equation describes a flux integral of the effective gravitational momentum across a surface of constant r. Taking this r to be near the horizon radius recovers the well-known result that the mass evolution of the black hole is directly related to the effective energy flux across the horizon associated with the timelike killing vector.

5.3.5.2 Evolution of full Kerr parameters

The computation to determine the first order mass and spin evolution of the full Kerr scenario proceeds similarly to the simpler Schwarzschild case explored in section 5.3.5.1. In the Kerr case, we will once again use the motivation of the close analogy to fluxes through the horizon and examine the quasistatic, spherically averaged part of the Ricci tensor. In Kerr, we find the relevant components for fixing the mass and spin evolution are R_{ww}, R_{wr}, $R_{r\phi}$, and $R_{w\phi}$.

We impose the analogous ansatz to the Schwarzschild case, but with the inclusion of slow spin variation and the understanding that the leading order metric to be varied $g^{(0)}_{\alpha\beta}$ is the Kerr metric in Boyer-Lindquist coordinates (5.5),

$$h^{(1)}_{\alpha\beta} = \frac{\partial g^{(0)}_{\alpha\beta}}{\partial M}\delta M^{(1)}(\tilde{w}) + \frac{\partial g^{(0)}_{\alpha\beta}}{\partial (aM)}\delta (aM)^{(1)}(\tilde{w}) + \mathcal{F}^{(1)}_{\alpha\beta}[\bar{x}^i, q^A, P^{(0)M}(\tilde{w})], \tag{5.48a}$$

$$h^{(2)}_{\alpha\beta} = \frac{\partial g^{(0)}_{\alpha\beta}}{\partial M}\delta M^{(2)}(\tilde{w}) + \frac{\partial g^{(0)}_{\alpha\beta}}{\partial (aM)}\delta (aM)^{(2)}(\tilde{w}) + \mathcal{F}^{(2)}_{\alpha\beta}[x^i, q^A, P^{(0)M}(\tilde{w})], \tag{5.48b}$$

As in the Schwarzschild case, we impose a gauge in which the $l = 0$, quasistatic part of the contribution $\mathcal{F}$ vanishes within the innermost point of the orbit $r < r^-$, for both first and second order perturbations,

$$\int d^3q^A \int d^2\Omega \delta R^{(1)}_{\alpha\beta}[\mathcal{F}^{(1)}]\Big|_{r<r^-} = 0, \tag{5.49a}$$

$$\int d^3q^A \int d^2\Omega \delta R^{(0)}_{\alpha\beta}[\mathcal{F}^{(2)}]\Big|_{r<r^-} = 0, \tag{5.49b}$$

$$\int d^3q^A \int d^2\Omega \delta R^{(0)}_{\alpha\beta}[\mathcal{F}^{(1)}]\Big|_{r<r^-} = 0. \tag{5.49c}$$

As the metric perturbation we are interested in constraining depends only on the slow-time parameter $\tilde{t}$, we take a fast-time average of the $l = 0$ component of the equation to determine the

slow evolution of the mass and spin. The relevant contributions to the Ricci tensor at $r < r^-$ are,

$$\int d^3q^A \int d^2\Omega R^{(1)l=0}_{rt}[h^{(1)}]\Big|_{r<r^-} = \frac{\pi\left(a^2 + r(3r-4M)\right)}{\Delta^2}\partial_{\tilde{t}}\delta M + \frac{\pi(M-r)(a^2+(2M-r)r)}{a\Delta^2}\partial_{\tilde{t}}\delta a \tag{5.50a}$$

$$\int d^3q^A \int d^2\Omega R^{(1)l=0}_{r\phi}[h^{(1)}]\Big|_{r<r^-} = -\frac{\pi(a^2-r^2)(a^2+r(r-4M))}{a\Delta^2}\partial_{\tilde{t}}\delta M - \frac{\pi M(3a^4-2a^2Mr+r^3(r-2M))}{a^2\Delta^2}\partial_{\tilde{t}}\delta a \tag{5.50b}$$

Consider once again the second order Einstein field equations (5.46), averaged over fast time and integrated over the sphere. The R_{ww}, R_{wr}, $R_{r\phi}$, and $R_{w\phi}$ components can then be inverted for δM and δa, which yields,

$$\begin{aligned}\partial_{\tilde{t}}\delta M =& \frac{M\left(-a^2+(2M-r)r\right)\left(-3a^4+2a^2Mr+(2M-r)r^3\right)}{\pi\left(r^4(r-2M)-2ra^2(r-2M)^2+a^4(r+2M)\right)} \\ &\times \int d^3q^A \int d^2\Omega R_{wr}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}] + h'(r)R_{ww}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}] \\ &+ \frac{a(M-r)(a^4-r^2(r-2M)^2)}{\pi(r^4(r-2M)-2ra^2(r-2M)^2+a^4(r+2M))} \\ &\times \int d^3q^A \int d^2\Omega R_{r\phi}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}] + h'(r)R_{w\phi}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}]\end{aligned} \tag{5.51a}$$

$$\begin{aligned}\partial_{\tilde{t}}\delta a =& \frac{-a(a-r)(a+r)\left(a^2+r(r-4M)\right)\left(a^2+r(r-2M)\right)}{\pi\left(r^4(r-2M)-2ra^2(r-2M)^2+a^4(r+2M)\right)} \\ &\times \int d^3q^A \int d^2\Omega R_{wr}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}] + h'(r)R_{ww}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}] \\ &+ \frac{-a^2(a^2+r(r-2M))\left(a^2+r(3r-4M)\right)}{\pi\left(r^4(r-2M)-2ra^2(r-2M)^2+a^4(r+2M)\right)} \\ &\times \int d^3q^A \int d^2\Omega R_{r\phi}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}] + h'(r)R_{w\phi}[\mathcal{F}^{(1)},\mathcal{F}^{(1)}].\end{aligned} \tag{5.51b}$$

5.4 Flux balance through second order for worldline frequencies

5.4.1 Motivation and similar identities

As the quasi-conserved orbital parameters evolve, a corresponding flux of energy and angular momentum should escape the system. The "flux balance" laws at first order in the mass ratio prove

this relation concretely and provide a means of computing dissipative dynamics via asymptotic gravitational wave amplitudes. The flux identities for first order perturbations have been previously derived from a number of strategies [371–373].

The flux balance laws offer the tantalizing possibility of significant savings in dissipative self force computations. The full form of the self force depends explicitly on the full metric perturbation. However, the rate of change of the quasi-conserved geodesic parameters can be re-cast from the self force expression to a form which depends only on scalar wave modes. Then, one can avoid computation of the full metric perturbation for dissipative quantities. Instead, one could compute only the Teukolsky modes, infer the instantaneous change of orbit parameters, and evolve the orbit accordingly.

To meet the sub-radian accuracy goal suggested in Section I, we will not be able to make the simplification of using dissipative fluxes at first order, as the first order conservative effects will produce relevant frequency shifts. However, as we discussed in Section II-B, the post-adiabatic waveform will require only the dissipative contributions at second order. We therefore develop an analogous flux balance identity for the second order metric perturbations.

The second order flux balance law is more intricate that the first order version. At second order, the nonlinear gravitational source contains a comparable effective energy and angular momentum to the amount that escapes the interaction zone. The effect of trapping this additional energy and angular momentum contribution has been previously discussed in the Post-Newtonian context [374], in which it is referred to as the Schott contribution, for its analogy to electromagnetic systems [375]. We find a flux balance equation at second order which can be described similarly: the energy lost by the orbit can be described as a sum over modes added to a volume integral over a subregion of the interaction zone.

In the two following subsections we derive the form of the flux balance laws by writing the metric perturbations as functionals of the full worldline z, motivated from the self-consistent construction [333], in which the self force takes comparatively simple form. The justification of this strategy can be found from a version of the self-consistent construction compatible with a multiscale expansion, which is discussed in the appendix of [344]. The version of these flux expressions in terms of the explicitly expanded, slowly varying worldline is given in Appendix 5.B.

5.4.2 First order evolution of E and L_z

To clarify the techniques that we will use to derive the flux balance law at second order, we first apply the strategy to the first order balance law and reproduce the previous results of [371, 373]. Our derivation strategy is closely modeled on the techniques used in [376].

For any killing vector of the background metric, $\nabla_{(\alpha}\xi_{\beta)} = 0$, the evolution of the geodesic parameters of the worldline is:

$$\frac{d\mathcal{E}}{d\tau} = u^\alpha \nabla_\alpha (u^\beta \xi_\beta) = a^\beta \xi_\beta + u^\alpha u^\beta \nabla_\alpha \xi_\beta = a^\beta \xi_\beta \tag{5.52}$$

We are interested in the orbit-averaged derivative of the quasi-conserved orbit parameters, so we perturbatively expand the orbit-averaged (5.52),

$$\varepsilon \left\langle \frac{d\mathcal{E}^{(0)}}{d\tilde{\tau}} \right\rangle + \varepsilon^2 \left\langle \frac{d\mathcal{E}^{(1)}}{d\tilde{\tau}} \right\rangle + \mathcal{O}(\varepsilon^2) = \varepsilon \left\langle a^{(1)\beta} \xi_\beta \right\rangle + \varepsilon^2 \left\langle a^{(2)\beta} \xi_\beta \right\rangle + \mathcal{O}(\varepsilon^2) \tag{5.53}$$

For convenience, the slow time dependence is parameterized as a dependence on $\tilde{\tau}(\tilde{w})$. As the left-hand side of this equation depends only on the first derivative, the appropriate worldline quantities required to relate the rate of energy change to $\langle d\mathcal{E}/d\tilde{w} \rangle$ may be inserted at the end of the computation.

Using the MiSaTaQuWa expression for the first order self force [377, 378], we expand the slow time derivative of the quasi-conserved orbit parameter as,

$$\begin{aligned}
\left\langle \frac{d\mathcal{E}^{(0)}}{d\tilde{\tau}} \right\rangle &= \left\langle \frac{-1}{2} (\xi^\beta + \xi_\alpha u^\alpha u^\beta)(2u^\delta u^\gamma \nabla_\delta h^{(1)\mathcal{R}}_{\gamma\beta} - u^\gamma u^\delta \nabla_\beta h^{(1)\mathcal{R}}_{\gamma\delta}) \right\rangle \\
&= \left\langle \frac{1}{2} \xi^\beta u^\gamma(\tau) u^\delta(\tau) \left(\nabla_\beta h^{(1)\mathcal{R}}_{\gamma\delta}(x)|_{x\to z(\tau)} \right) \right\rangle - \left\langle (\xi^\beta + \frac{1}{2}\mathcal{E} u^\beta) u^\gamma u^\delta \left(\nabla_\delta h^{(1)\mathcal{R}}_{\beta\gamma}(x)|_{x\to z(\tau)} \right) \right\rangle .
\end{aligned} \tag{5.54}$$

At this point, we simplify by commuting terms through the derivatives. During this process, we discard any value which depends on the acceleration a^μ, as any such term is quadratic in the first order metric perturbation, and therefore a second order contribution. In the next section, we will restore each of these residual terms to appropriately derive the second-order evolution of the

quasi-conserved quantities.

$$\begin{aligned}\left\langle \frac{d\mathcal{E}}{d\tilde{\tau}} \right\rangle =& \left\langle \frac{1}{2}(u^\gamma u^\delta \partial_\xi h^{(1)\mathcal{R}}_{\gamma\delta}) \right\rangle - \left\langle \Gamma^\gamma_{\lambda\eta} \xi^\eta u^\gamma h^{(1)\mathcal{R}}_{\gamma\delta} u^\delta \right\rangle - \left\langle \xi^\beta u^\gamma u^\delta \nabla_\delta h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle - \left\langle \frac{1}{2}\mathcal{E} u^\beta u^\gamma u^\delta \nabla_\delta h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle \\ =& \left\langle \frac{1}{2}(u^\gamma u^\delta \partial_\xi h^{(1)\mathcal{R}}_{\gamma\delta}) \right\rangle - \left\langle u^\gamma \nabla_\gamma (\xi^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma}) \right\rangle - \left\langle \frac{1}{2} u^\delta \nabla_\delta (\mathcal{E} u^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma}) \right\rangle . \end{aligned} \tag{5.55}$$

When performing the average over fast-time modes q^A, the total derivatives with respect to proper time pick out the non-zero frequency parts, which then average to zero. We are left with only the first term:

$$\left\langle \frac{d\mathcal{E}}{d\tau} \right\rangle = \frac{1}{2} \left\langle u^\alpha u^\beta \partial_\xi h^{(1)\mathcal{R}}_{\alpha\beta} \right\rangle . \tag{5.56}$$

The metric perturbation which appears in the equation (5.56) is the regular metric perturbation. The flux formula (5.56) has an important symmetry property for the reciprocity of the Green's functions. To examine this property, we expand the implicit Green's function integral which is used to derive the first order regular field,

$$\frac{1}{2}\langle u^\alpha u^\beta \mathcal{L}_\xi h^{(1)\mathcal{R}}_{\alpha\beta} \rangle = \int d^3q^A \int d^3q'^A u^\alpha u^\beta \mathcal{L}_\xi G^{\mathcal{R}}_{\alpha\beta\gamma'\delta'}(z(q^A), z(q'^B) u^{\gamma'} u^{\delta'} . \tag{5.57}$$

However, the Green's function has the reciprocity property $\mathcal{L}_\xi G(x,x') = -\mathcal{L}_{\xi'} G(x,x')$, so the symmetry of the integrals above allow us to re-write the flux expression as

$$\begin{aligned}\frac{1}{2}\langle u^\alpha u^\beta \mathcal{L}_\xi h^{(1)\mathcal{R}}_{\alpha\beta} \rangle =& \int d^3q^A \int d^3q'^A u^\alpha u^\beta \left(\mathcal{L}_\xi G^{\mathcal{R}}_{\alpha\beta\gamma'\delta'}(z(q^A), z(q'^B) - \mathcal{L}_\xi G^{\mathcal{R}}_{\alpha\beta\gamma'\delta'}(z(q^A), z(q'^B)) \right) u^{\gamma'} u^{\delta'} \\ =& \frac{1}{2}\langle u^\alpha u^\beta \mathcal{L}_\xi h^{(1)\mathrm{Rad}}_{\alpha\beta} \rangle, \end{aligned} \tag{5.58}$$

where the $h^{(1)\mathrm{Rad}}_{\alpha\beta}$ denotes the radiative part of the metric perturbation, obtained from the half-retarded minus half-advanced Green's functions.

To complete the association with asymptotic fluxes, consider the mode decomposition of the radiative solution $h^{(1)\mathrm{Rad}}_{\alpha\beta}$ [357, 358]:

$$\begin{aligned} h^{(1)\mathrm{Rad}}_{\alpha\beta} =& \frac{1}{8\pi i} \int_{-\infty}^{\infty} d\omega \frac{\omega}{|\omega|} \sum_{l=2}^{\infty} \sum_{m=-l}^{l} \left[\frac{1}{\alpha^*_{-s\omega lm} \alpha_{s\omega lm}} {}_sZ^{(1)\mathrm{out}}_{\omega lm} \; {}_s\tau^\dagger_{\alpha\beta}(x) {}_s\Phi^{\mathrm{out}}_{\omega lm} \right. \\ & \left. + \frac{\kappa_{s\omega m} \tau_{s\omega lm} \tau^*_{-s\omega lm}}{\beta_{s\omega lm} \beta^*_{-s\omega lm}} \frac{\omega p_{m\omega}}{|\omega p_{m\omega}|} {}_sZ^{(1)\mathrm{down}}_{\omega lm} \; {}_s\tau^\dagger_{\alpha\beta} \; {}_s\Phi^{\mathrm{down}}_{\omega lm} \right] \\ \equiv& \frac{1}{8\pi i} \int_{-\infty}^{\infty} d\omega \sum_{l=2}^{\infty} \sum_{m=-l}^{l} \left[\mathcal{A}^{\mathrm{out}}_{s\omega lm} \; {}_sZ^{(1)\mathrm{out}}_{\omega lm} \; {}_s\tau^\dagger_{\alpha\beta}(x) {}_s\Phi^{\mathrm{out}}_{\omega lm} + \mathcal{A}^{\mathrm{down}}_{s\omega lm} {}_sZ^{(1)\mathrm{down}}_{\omega lm} \; {}_s\tau^\dagger_{\alpha\beta} \; {}_s\Phi^{\mathrm{down}}_{\omega lm} \right], \end{aligned} \tag{5.59}$$

where the radiative out and down mode amplitudes are defined as,

$$ {}_sZ^{\rm out}_{\omega lm} = \alpha^*_{-s\omega lm}\alpha_{s\omega lm}\int d^4x\sqrt{-g}\ {}_s\Phi^{\rm down}_{\omega lm}\ {}_s\tau_{\alpha\beta}T^{\alpha\beta} \tag{5.60a} $$

$$ {}_sZ^{\rm down}_{\omega lm} = \frac{\beta^*_{-s\omega lm}\beta_{s\omega lm}}{\kappa_{s\omega m}\tau_{s\omega lm}\tau^*_{-s\omega lm}}\frac{\omega p_{m\omega}}{|\omega p_{m\omega}|}\int d^4x\sqrt{-g}\ {}_s\Phi^{\rm down}_{\omega lm}\ {}_s\tau_{\alpha\beta}T^{\alpha\beta}. \tag{5.60b} $$

and the functions $\Phi^{\rm out/down}$ are the separable modes of the Hertz potential, described in Section 5.3.3.

Finally, expanding (5.56) for the timelike killing vector and performing the integration over fast time variables q^A, we find

$$ \left\langle \frac{dE}{d\tilde{\tau}} \right\rangle = \frac{1}{2}\int_{-\infty}^{\infty} i\omega \sum_{l=2}^{\infty}\sum_{m=-l}^{l} \frac{\mathcal{A}^{*\,\rm out}_{s\omega lm}\mathcal{A}^{\rm out}_{s\omega lm}}{\mu}\left(Z_V^{(1)\rm out}\right)^2 + \frac{\mathcal{A}^{*\,\rm down}_{s\omega lm}\mathcal{A}^{\rm down}_{s\omega lm}}{\mu}\left(Z_V^{(1)\rm down}\right)^2, \tag{5.61} $$

which are amplitude values computable entirely from the amplitudes of the $\Phi^{\rm out}$ and $\Phi^{\rm down}$ modes distant from the inspiral.

Similarly, the first derivative of the orbital angular momentum can be evaluated as

$$ \left\langle \frac{dL_z}{d\tau} \right\rangle = \frac{1}{2}\int_{-\infty}^{\infty} im \sum_{l=2}^{\infty}\sum_{m=-l}^{l} \frac{\mathcal{A}^{*\,\rm out}_{s\omega lm}\mathcal{A}^{\rm out}_{s\omega lm}}{\mu}\left(Z_V^{(1)\rm out}\right)^2 + \frac{\mathcal{A}^{*\,\rm down}_{s\omega lm}\mathcal{A}^{\rm down}_{s\omega lm}}{\mu}\left(Z_V^{(1)\rm down}\right)^2. \tag{5.62} $$

5.4.3 Second order evolution of E and L_z

The second order dissipative effects will include both the expansion of the second order self acceleration, analogous to computation performed in the previous subsection, and the inclusion of all $\mathcal{O}(\varepsilon^2)$ acceleration terms which were neglected from the first order computation to bring it to the compact form (5.56). We refer to these additional acceleration terms from the previous computation as the residual self-acceleration. We write the dissipative second order evolution as

$$ \left\langle \frac{d\mathcal{E}^{(1)}}{d\tilde{\tau}} \right\rangle = \left\langle \xi_\alpha a^{(1)\alpha}[h^{(2)}] \right\rangle + \left\langle \xi_\alpha a^{(2)\alpha}[h^{(1)}, h^{(1)}] \right\rangle + \left\langle \left(\frac{d\mathcal{E}^{(0)}}{d\tilde{\tau}}\right)^{(1)}_{\rm res} \right\rangle. \tag{5.63} $$

The residual from the first-order alterations is found by maintaining all factors of the acceleration when commuting u^μ through the covariant derivatives in the computation performed in equations (5.54,5.55). The residual term which results from that expansion is then,

$$ \left\langle \left(\frac{d\mathcal{E}}{d\tilde{\tau}}\right)^{(1)}_{\rm res.} \right\rangle = \left\langle \xi^\beta a^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle + \frac{1}{2}\left\langle u^\delta \nabla_\delta \mathcal{E} u^\gamma u^\beta h^{(1)\mathcal{R}}_{\gamma\beta} \right\rangle + \left\langle \mathcal{E} a^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle. \tag{5.64} $$

Each of the values in this residual must be expanded using the MiSaTaQuWa self force expression to a quadratic expression of $h^{(1)\mathcal{R}}$ (removing immediately total derivatives with respect to τ and terms of order $\mathcal{O}(\varepsilon^3)$ or higher).

$$\left\langle \xi^\beta a^{(1)\gamma}[h^{(1)\mathcal{R}}] h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle = -\left\langle \xi_\delta h^{(1)\mathcal{R}\delta\beta} u^\gamma u^\lambda \nabla_\lambda h^{(1)\mathcal{R}}_{\gamma\beta} \right\rangle + \frac{1}{2}\left\langle \xi_\delta h^{(1)\mathcal{R}\delta\beta} u^\gamma u^\lambda \nabla_\beta h^{(1)\mathcal{R}}_{\gamma\lambda} \right\rangle - \frac{1}{2}\left\langle \xi_\delta h^{(1)\mathcal{R}\delta\beta} u_\beta u^\eta u^\gamma u^\lambda \nabla_\lambda h^{(1)\mathcal{R}}_{\gamma\eta} \right\rangle \tag{5.65}$$

$$\left\langle \frac{1}{2}\frac{d\mathcal{E}^{(0)}}{d\tilde{\tau}} u^\gamma u^\beta h^{(1)\mathcal{R}}_{\gamma\beta} \right\rangle = \frac{1}{8}\left\langle u^\alpha u^\beta u^\gamma u^\delta \partial_\xi (h^{(1)\mathcal{R}}_{\alpha\beta} h^{(1)\mathcal{R}}_{\gamma\delta}) \right\rangle + \frac{1}{2}\left\langle \xi^\gamma u^\delta h^{(1)\mathcal{R}}_{\gamma\delta} u^\alpha u^\lambda u^\beta \nabla_\beta h^{(1)\mathcal{R}}_{\alpha\lambda} \right\rangle \tag{5.66}$$

$$\left\langle \mathcal{E} a^{(1)\beta} u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle = \frac{1}{2}\left\langle \mathcal{E} u_\delta h^{(1)\mathcal{R}\delta\gamma} u^\eta u^\sigma \nabla_\gamma h^{(1)\mathcal{R}}_{\eta\sigma} \right\rangle \tag{5.67}$$

The second-order acceleration term can be expanded using [351, 352] to give (removing total derivatives with respect to τ):

$$\begin{aligned}\left\langle \xi_\alpha a^{(2)\alpha}[h^{(1)\mathcal{R}}, h^{(1)\mathcal{R}}] \right\rangle =& \frac{1}{2}\left\langle \xi_\mu (g^{\mu\nu} + u^\mu u^\nu)(-h^{(1)\mathcal{R}\rho}_\nu) u^\sigma u^\lambda (\nabla_\rho h^{(1)\mathcal{R}}_{\sigma\lambda} - 2\nabla_\lambda h^{(1)\mathcal{R}}_{\sigma\rho}) \right\rangle \\ =& -\frac{1}{2}\left\langle \xi_\nu h^{(1)\mathcal{R}\nu\rho} u^\sigma u^\lambda \nabla_\rho h^{(1)\mathcal{R}}_{\sigma\lambda} \right\rangle - \frac{1}{2}\left\langle \mathcal{E} u_\nu h^{(1)\mathcal{R}\nu\rho} u^\sigma u^\lambda \nabla_\rho h^{(1)\mathcal{R}}_{\sigma\lambda} \right\rangle \\ &+ \left\langle \xi_\nu h^{(1)\mathcal{R}\nu\rho} u^\sigma u^\lambda \nabla_\lambda h^{(1)\mathcal{R}}_{\sigma\rho} \right\rangle \end{aligned} \tag{5.68}$$

However, in these terms, we've neglected the slow-time derivatives that should be included from the terms canceled as orbit-averages of total time derivatives at first order. In particular, those slow variations give additional contributions,

$$-\left\langle \partial_{\tilde{\tau}} \left(\xi^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right) \right\rangle - \frac{1}{2}\left\langle \partial_{\tilde{\tau}} \left(\mathcal{E} u^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right) \right\rangle \tag{5.69}$$

Summing the various contributions to (5.63), we obtain the compact final expression for the orbit parameter evolution:

$$\begin{aligned}\left\langle \frac{d\mathcal{E}^{(1)}}{d\tilde{\tau}} \right\rangle =& \frac{1}{2}\left\langle u^\alpha u^\gamma \partial_\xi h^{(2)\mathcal{R}}_{\alpha\gamma} \right\rangle + \frac{1}{8}\left\langle u^\alpha u^\beta u^\gamma u^\delta \partial_\xi (h^{(1)\mathcal{R}}_{\alpha\beta} h^{(1)\mathcal{R}}_{\gamma\delta}) \right\rangle \\ &- \partial_{\tilde{\tau}}\left\langle \xi^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle - \frac{1}{2}\partial_{\tilde{\tau}}\left\langle \mathcal{E} u^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle . \end{aligned} \tag{5.70}$$

This expression for the orbit evolution is gauge invariant under smooth gauge transformations consistent with the multiscale construction. The gauge invariance is demonstrated in appendix 5.B.

The average energy and angular momentum loss can be evaluated using only full metric information at first order and the amplitudes of the Teukolsky modes at second order far from the

inspiraling system. This statement is the analogous second-order form to the 'balance law' at first order: as the inspiraling body radiates conserved energy or angular momentum, a well-defined energy or angular momentum leaves the system, and the difference between these two quantities is expressed as a 'Schott' [375] term associated with an amount of energy or angular momentum that is retained within the bulk of the system. In the next several steps of this computation, we will process the dissipation formula (5.70) to a formal expression for a second-order flux balance law.

The regular second order metric perturbation is not a homogeneous solution. Effective source methods demonstrate that there exists a regularized, smooth source function $S_{(\mathrm{Eff})\,\mu\nu}[h^{(1)}, h^{(2)}, z^\mu]$, which can be formally written

$$S_{(\mathrm{Eff})\,\mu\nu}[h^{(1)}, h^{(2)}, z^\mu] = R_{\mu\nu}[h^{(1)}, h^{(1)}] - E_{\mu\nu}[h^{(2)\mathcal{P}}], \tag{5.71}$$

where $h^{(2)\mathcal{P}}$ is the puncture field constructed by [Pound 2017], and which is equivalent to the non-distributional quadratic Ricci source arising from the singular components $h^{(1)S}$. The decomposition of the first order metric gives the first order metric perturbation within the worldtube puncture region as $h^{(1)}|_T = h^{(1)\mathcal{R}} + h^{(1)\mathcal{P}}$, where $h^{(1)\mathcal{R}}$ is a homogeneous solution, and the metric perturbation outside the worldtube puncture region $h^{(1)}|_{\mathcal{M}\setminus T}$ is a homogeneous solution. We therefore introduce the expression $\int d^4x S^{\alpha\beta}_{\mathrm{eff}} \partial_\xi h^{(1)}_{\alpha\beta}$ to the dissipation formula (5.70) in order to make use of the Green's function reciprocity relations,

$$\begin{aligned}\left\langle \frac{d\mathcal{E}^{(1)}}{d\tilde{\tau}} \right\rangle =& \frac{1}{2}\left(\left\langle u^\alpha u^\beta \partial_\xi h^{(2)\mathcal{R}}_{\alpha\beta} \right\rangle + \int d^4x S^{\alpha\beta}_{\mathrm{eff}} \partial_\xi h^{(1)}_{\alpha\beta} \right) - \frac{1}{2\mu}\int_{\mathcal{M}\setminus T} d^4x \sqrt{-g} S^{\alpha\beta}_{\mathrm{eff}} \partial_\xi h^{(1)}_{\alpha\beta} \\ &- \frac{1}{2\mu}\int_T d^4x \sqrt{-g} S^{\alpha\beta}_{\mathrm{eff}} \partial_\xi h^{(1)\mathcal{R}}_{\alpha\beta} - \frac{1}{2\mu}\int_T d^4x \sqrt{-g} S^{\alpha\beta}_{\mathrm{eff}} \partial_\xi h^{(1)\mathcal{P}}_{\alpha\beta} \\ &+ \frac{1}{8}\left\langle u^\alpha u^\beta u^\gamma u^\delta \partial_\xi (h^{(1)\mathcal{R}}_{\alpha\beta} h^{(1)\mathcal{R}}_{\gamma\delta}) \right\rangle - \partial_{\tilde{\tau}} \left\langle \xi^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle - \frac{1}{2}\partial_{\tilde{\tau}} \left\langle \mathcal{E} u^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle, \end{aligned} \tag{5.72}$$

where the integration regions T and $\mathcal{M}\setminus T$ denote the worldtube and the full manifold outside the worldtube, respectively.

We develop a procedure for computing the above energy and angular momentum loss formula in terms of quantities that can be computed as asymptotic fluxes of Teukolsky variables and first order full metric perturbations alone. We also attempt to confine any integration which is necessary to compute these quantities to sub-regions of the spacetime to provide efficiency gains over simply computing the second order field in entirety. The second line of (5.72) satisfies these requirements.

All integrals involved are integrals only over the worldline, and only over first order regular fields. There are several contributions in the first line which seem comparatively difficult to compute, but we demonstrate that many of them can be re-written in terms of flux expressions.

Consider first the pair of terms

$$\begin{aligned}\left\langle u^\alpha u^\beta \partial_\xi h^{(2)\mathcal{R}}_{\alpha\beta}\right\rangle &+ \int d^4x S^{\alpha\beta}_{\text{eff}} \partial_\xi h^{(1)}_{\alpha\beta} \\ &= \int d\tau \int d^4x u^\alpha u^\beta \partial_\xi G^{\text{ret}}_{\alpha\beta\alpha'\beta'}(z(\tau),x) S^{\alpha'\beta'}_{\text{eff}}(x) + S^{\alpha'\beta'}_{\text{eff}}(x)\partial_{\xi'} G^{\text{ret}}_{\alpha'\beta'\alpha\beta}(x,z(\tau)) u^\alpha u^\beta \\ &= \int d\tau \int d^4x u^\alpha u^\beta \left(\partial_\xi G^{\text{ret}}_{\alpha\beta\alpha'\beta'}(z(\tau),x) - \partial_\xi G^{\text{ret}}_{\alpha'\beta'\alpha\beta}(x,z(\tau))\right) S^{\alpha'\beta'}_{\text{eff}}(x) \\ &= \int d\tau \int d^4x u^\alpha u^\beta \partial_\xi G^{\text{Rad}}_{\alpha\beta\alpha'\beta'}(z(\tau),x) S^{\alpha'\beta'}_{\text{eff}}(x) \\ &= \left\langle u^\alpha u^\beta \partial_\xi h^{(2)\text{Rad}}_{\alpha\beta}\right\rangle \end{aligned} \tag{5.73}$$

The radiative part of the second order metric perturbation may be re-written, up to a gauge transformation, as the sum over reconstructed modes,

$$h^{(2)\text{Rad}}_{\alpha\beta} = \frac{1}{8\pi i}\int_{-\infty}^{\infty} d\omega \sum_{\{l,m,s\}} \mathcal{A}_{s\omega lms} Z^{(2)\text{ out}}_{\omega lm}\, {}_s\tau^\dagger_{\alpha\beta}\, {}_s\Phi^{\text{out}}_{\omega lm} + \mathcal{B}_{s\omega lms} Z^{(2)\text{ down}}_{\omega lm}\, {}_s\tau^\dagger_{\alpha\beta}\, {}_s\Phi^{\text{down}}_{\omega lm}, \tag{5.74}$$

where the second order mode amplitudes are given directly by evaluating the radiative part of the Teukolsky modes. The radiative modes may be derived via [357, 358],

$$\begin{aligned} {}_sZ^{\text{down}}_{\omega lm} =& \frac{1}{{}_s\mathcal{A}_{\omega lm}} \int d^4x' \sqrt{-g(x')}\, {}_s\Phi^{\text{down}}_{\omega lm}(x')^*\, {}_s\tau_{\alpha\beta} S^{\alpha\beta}_{\text{eff}}(x'), \\ =& \frac{1}{{}_s\mathcal{A}_{\omega lm}} \int d^4x' \sqrt{-g(x')} \left({}_s\tau^\dagger_{\alpha\beta}\, {}_s\Phi^{\text{down}}_{\omega lm}(x')\right)^* S^{\alpha\beta}_{\text{eff}}(x'), \end{aligned} \tag{5.75}$$

and

$$\begin{aligned} {}_sZ^{\text{out}}_{\omega lm} =& \frac{1}{{}_s\mathcal{B}_{\omega lm}} \int d^4x' \sqrt{-g(x')}\, {}_s\Phi^{\text{out}}_{\omega lm}(x')^*\, {}_s\tau_{\alpha\beta} S^{\alpha\beta}_{\text{eff}}(x'), \\ =& \frac{1}{{}_s\mathcal{B}_{\omega lm}} \int d^4x' \sqrt{-g(x')} \left({}_s\tau^\dagger_{\alpha\beta}\, {}_s\Phi^{\text{out}}_{\omega lm}(x')\right)^* S^{\alpha\beta}_{\text{eff}}(x'). \end{aligned} \tag{5.76}$$

Therefore, we re-write the result from (5.73), up to a gauge transformation, as a sum over radiative modes,

$$\left\langle u^\alpha u^\beta \partial_\xi h^{(2)\text{Rad}}_{\alpha\beta}\right\rangle = \frac{1}{8\pi i}\int_{-\infty}^{\infty} d\omega \sum_{l,m,s} \frac{\mathcal{A}^2_{s\omega lm}}{\mu}\, {}_sZ^{(2)\text{out}}_{\omega lm}\, {}_sZ^{(1)\text{out}}_{\omega lm} + \frac{\mathcal{B}^2_{s\omega lm}}{\mu}\, {}_sZ^{(2)\text{down}}_{\omega lm}\, {}_sZ^{(1)\text{down}}_{\omega lm} \tag{5.77}$$

Consider now the third term of the first line of (5.72). We note that the first-order metric perturbation away from the worldline is a homogeneous solution to the relaxed EFE, and therefore also possesses a mode decomposition in the radiation gauge. It is, however, notable that such mode amplitudes will not be given by the radiative formulas (5.75), (5.76). Instead, I will denote the resulting mode amplitudes by $\tilde{Z}^{(1)\text{out/down}}$:

$$h^{(1)}_{\alpha\beta}\big|_{\mathcal{M}\setminus T} = h^{(1)\text{comp}}_{\alpha\beta} + \frac{1}{8\pi i}\int_{-\infty}^{\infty} d\omega \sum_{\{l,m,s\}} \mathcal{A}_{s\omega lms}\tilde{Z}^{(1)\text{ out}}_{\omega lm}\, {}_s\tau^{\dagger}_{\alpha\beta}\, {}_s\Phi^{\text{out}}_{\omega lm} + \mathcal{B}_{s\omega lms}\tilde{Z}^{(1)\text{ down}}_{\omega lm}\, {}_s\tau^{\dagger}_{\alpha\beta}\, {}_s\Phi^{\text{down}}_{\omega lm}, \tag{5.78}$$

where $h^{(1)\text{comp}}_{\alpha\beta}$ denotes the $l = 0$, $l = 1$ completion part of the metric perturbation, comprising gauge modes, mass, and angular momentum content. Similarly, the regular solution $h^{(1)\mathcal{R}}$ is a homogeneous solution in the neighborhood of the worldline, and permits a mode decomposition when evaluated within the worldtube.

$$h^{(1)\mathcal{R}}_{\alpha\beta}\big|_{T} = h^{(1)\mathcal{R}\,\text{comp}}_{\alpha\beta} + \frac{1}{8\pi i}\int_{-\infty}^{\infty} d\omega \sum_{\{l,m,s\}} \mathcal{A}_{s\omega lms}\tilde{Z}^{(1)\mathcal{R}\text{ out}}_{\omega lm}\, {}_s\tau^{\dagger}_{\alpha\beta}\, {}_s\Phi^{\text{out}}_{\omega lm} + \mathcal{B}_{s\omega lms}\tilde{Z}^{(1)\mathcal{R}\text{ down}}_{\omega lm}\, {}_s\tau^{\dagger}_{\alpha\beta}\, {}_s\Phi^{\text{down}}_{\omega lm}, \tag{5.79}$$

The third term of the first line of (5.72) may then be expanded as

$$-\frac{1}{2}\int_{\mathcal{M}\setminus T} d^4x\sqrt{-g}S^{\alpha\beta}_{\text{eff}}\partial_\xi h^{(1)}_{\alpha\beta} = \frac{-1}{16\pi i}\int_{-\infty}^{\infty} d\omega \sum_{\{l,m,s\}} \frac{\mathcal{A}^2_{s\omega lm}}{\mu}\, {}_s\mathcal{Z}^{(2)\mathcal{M}\setminus T\,\text{out}}_{\omega lm}\, {}_s\tilde{Z}^{(1)\text{out}}_{\omega lm} + \frac{\mathcal{B}^2_{s\omega lm}}{\mu}\, {}_s\mathcal{Z}^{(2)\mathcal{M}\setminus T\,\text{down}}_{\omega lm}\, {}_s\tilde{Z}^{(1)\text{down}}_{\omega lm}, \tag{5.80}$$

where we have introduced mode amplitudes $\mathcal{Z}^{(2)\mathcal{M}\setminus T\,\text{out/down}}$ associated with integration over the region which excludes the worldtube:

$$\begin{aligned} {}_s\mathcal{Z}^{(2)\mathcal{M}\setminus T\,\text{down}}_{\omega lm} &\equiv \frac{1}{{}_s\mathcal{A}_{\omega lm}}\int_{\mathcal{M}\setminus T} d^4x'\sqrt{-g(x')}\left({}_s\tau^{\dagger}_{\alpha\beta}\, {}_s\Phi^{\text{down}}_{\omega lm}(x')\right)^* S^{\alpha\beta}_{\text{eff}}(x'), \\ {}_s\mathcal{Z}^{(2)\mathcal{M}\setminus T\,\text{out}}_{\omega lm} &\equiv \frac{1}{{}_s\mathcal{B}_{\omega lm}}\int_{\mathcal{M}\setminus T} d^4x'\sqrt{-g(x')}\left({}_s\tau^{\dagger}_{\alpha\beta}\, {}_s\Phi^{\text{out}}_{\omega lm}(x')\right)^* S^{\alpha\beta}_{\text{eff}}(x'). \end{aligned} \tag{5.81}$$

Finally, the fourth term of the first line of (5.72) can be decomposed as

$$-\frac{1}{2}\int_{T} d^4x\sqrt{-g}S^{\alpha\beta}_{\text{eff}}\partial_\xi h^{(1)\mathcal{R}}_{\alpha\beta} = \frac{-1}{16\pi i}\int_{-\infty}^{\infty} d\omega \sum_{\{l,m,s\}} \frac{\mathcal{A}^2_{s\omega lm}}{\mu}\, {}_s\mathcal{Z}^{(2)T\,\text{out}}_{\omega lm}\, {}_s\tilde{Z}^{(1)\mathcal{R}\,\text{out}}_{\omega lm} + \frac{\mathcal{B}^2_{s\omega lm}}{\mu}\, {}_s\mathcal{Z}^{(2)T\,\text{down}}_{\omega lm}\, {}_s\tilde{Z}^{(1)\mathcal{R}\,\text{down}}_{\omega lm}, \tag{5.82}$$

The remaining terms of (5.72) may not be easily decomposed, but can be computed entirely from the metric perturbations in the vicinity of the worldline.

We must now consider the question of which gauge each of these quantities will be evaluated in. In order to use the first order flux decomposition into mode amplitudes, we have already committed to the radiation gauge in the vicinity of the worldline at first order. Additionally, each of the terms in the first line of (5.72) must be converted to the radiation gauge in order to support the mode decompositions.

As the entire expression is gauge-invariant under smooth gauge transformations, we might hope to put the entire expression in radiation gauge. However, this is computationally inconvenient, as several parts of the expansion (5.72) will either not support a radiation gauge representation (in the case of $h^{\mathcal{P}}$), or include contributions from the problematic completion part (in the case of the expressions dependent on $h^{\mathcal{R}}$).

We therefore include in the final flux formula explicit gauge-transformation terms acquired from the conversion of each of the required flux terms to radiation gauge. These terms are computed from intermediate steps in the Appendix section 5.B. Our final balance law formula is then,

$$\begin{aligned}
\left\langle \frac{\partial \mathcal{E}^{(1)}}{\partial \tilde{\tau}} \right\rangle =& \frac{1}{16\pi i} \int_{-\infty}^{\infty} d\omega \sum_{\{l,m,s\}} \frac{\mathcal{A}^2_{s\omega lm}}{\mu} \Big({}_sZ^{(2)\text{out}}_{\omega lm} \, {}_sZ^{(1)\text{out}}_{\omega lm} - {}_s\mathcal{Z}^{(2)\mathcal{M}\setminus T\,\text{out}}_{\omega lm} \, {}_s\tilde{Z}^{(1)\text{out}}_{\omega lm} \\
&\qquad - {}_s\mathcal{Z}^{(2)T\,\text{out}}_{\omega lm} \, {}_s\tilde{Z}^{(1)\mathcal{R}\,\text{out}}_{\omega lm} \Big) \\
&+ \frac{\mathcal{B}^2_{s\omega lm}}{\mu} \Big({}_sZ^{(2)\text{down}}_{\omega lm} \, {}_sZ^{(1)\text{down}}_{\omega lm} - {}_s\mathcal{Z}^{(2)\mathcal{M}\setminus T\,\text{down}}_{\omega lm} \, {}_s\tilde{Z}^{(1)\text{down}}_{\omega lm} \\
&\qquad - {}_s\mathcal{Z}^{(2)T\,\text{down}}_{\omega lm} \, {}_s\tilde{Z}^{(1)\mathcal{R}\,\text{down}}_{\omega lm} \Big) \\
&+ \frac{1}{8} \left\langle u^\alpha u^\beta u^\gamma u^\delta \partial_\xi (h^{(1)\mathcal{R}}_{\alpha\beta} h^{(1)\mathcal{R}}_{\gamma\delta}) \right\rangle - \partial_{\tilde{\tau}} \left\langle \xi^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle - \frac{1}{2} \partial_{\tilde{\tau}} \left\langle \mathcal{E} u^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma} \right\rangle \\
&- \frac{1}{2} \left\langle u^\mu u^\nu (\mathcal{L}_\zeta + \pounds_\zeta) \mathcal{L}_\xi h^{(1)\mathcal{R}}_{\mu\nu} \right\rangle - \frac{1}{2} \varepsilon^2 \left\langle u^\mu u^\nu \mathcal{L}_\zeta \mathcal{L}_\xi \mathcal{L}_\zeta g^{(0)}_{\mu\nu} \right\rangle \\
&- \frac{1}{4} \left\langle u^\alpha u^\beta \mathcal{L}_\zeta g^{(0)}_{\alpha\beta} u^\mu u^\nu \left(\mathcal{L}_\xi h^{(1)\mathcal{R}}_{\mu\nu} + \mathcal{L}_\xi \mathcal{L}_\zeta g^{(0)}_{\mu\nu} \right) \right\rangle + \partial_{\tilde{\tau}} \left\langle \xi^\beta u^\gamma (\nabla_\beta \zeta_\gamma + \nabla_\gamma \zeta_\beta) \right\rangle ,
\end{aligned} \tag{5.83}$$

where ζ is the vector necessary to transform the first order $h^{(1)\text{Rad}}_{\alpha\beta}$ from the Lorenz gauge to the radiation gauge.

5.5 Conclusions

We have presented the field equations necessary to complete the description of the multiscale approximation method introduced in [306], for the interaction region $|r^*| \ll M/\varepsilon$. We present the multiscale evolution equations for a post-adiabatic Teukolsky variable computation in Section 5.3, in which we develop the necessary second order source associated with the slow variation of the first order solution. In addition, the first order correction to the worldline must be included as a small dipole correction to the puncture metric used to regularize the second order source. Finally, the slow evolution of the large companion's mass and spin must be included, and is computed from consistency with the multiscale construction (for the TLC computation).

The full procedure for evolving the field equations in the interaction zone is accomplished as the sequence of steps, for each slow-time step:

1. Compute (or cache) the instantaneous leading trajectory of the small companion from the leading geodesic parameters (From Kerr Celestial mechanics, [306, 344] or Appendix 5.A)
2. Compute (or cache) the first order metric perturbation from the leading trajectory [344]
3. Use the first order metric perturbation to derive the slow-time evolution of the leading geodesic parameters [306, 344]
4. Use the first order metric perturbation to derive the slow-time evolution of the central mass and spin (Section 5.3.5.2)
5. Match the first order metric perturbation to Near-horizon zone and Far zone computations and compute the subleading corrections in near-horizon and far zones [345, 346]
6. Compute the correction to worldline puncture from the subleading geodesic parameters and first order self force
7. Using first order quadratic source, slow-time evolution of first order field, the subleading puncture correction, and the asymptotic matching information from near-horizon and far zones to compute the second-order metric perturbation (Section 5.3)

8. Either

   a) use the first and second order metric perturbation to compute the second-order self force, and find the slow evolution of the subleading orbit parameters from the angle-average [351,352],

   or

   b) se the first and second order metric perturbation to find the subleading flux to determine the slow evolution of the subleading orbit parameters. (Section 5.4)

9. Extract the waveform as a function of fast time at future null infinity from the Far zone solution [345]

10. Evolve the leading, subleading geodesic quantities and mass and spin corrections

11. Repeat for subsequent slow times

In addition to the multiscale field equations, we have presented the generalization of the flux-balance laws to second order. At first order, the flux balance laws give a direct equivalence between the energy lost by the slow inspiral and the flux of effective stress energy associated with the quadratic dependence of the Ricci tensor on the metric perturbation far from the inspiral. At second order, we derive an analogous, but more intricate equation. At second order, the energy lost by the orbit can either by dissipated to asymptotic fluxes or be trapped in the bulk of the manifold as Schott terms.

The remaining ingredients, which we give an overview for in Chapter 4, are the computation of the post-Minkowski approximation far from the system $r^* \gg M$, and the computation of the near-horizon approximation very close to the horizon of the large companion $-r^* \gg M$. The post-Minkowski approximation method will give the details of the evolution of waveforms from the interface with the interaction region to future null infinity, as well as developing additional quasistatic contributions needed as inputs to the integration suggested for the interaction region. The post-Minkowski and geometric optics approximations will be the subject of a forthcoming paper [345], and we discuss an overview of the post-adiabatic Far Zone calculation in 6. The near-horizon region also requires a separate tactic, and similarly provides valuable boundary information

as inputs to the second-order integration in the interaction region. The near-horizon approximation will be discussed in a separate forthcoming paper [346].

5.A | Details of the subleading worldine $z^{(1)}$

5.A.1 | Near-identity transformation of geodesic parameters

We would like to cleanly separate periodic dependence from the slow radiation-reaction dependence in the evolution of the inspiral system. This distinction is closely related to the separation between a dissipative and conservative self force, but we will find it necessary to make a somewhat different split between the two components than is traditionally considered.

We propose a new set of action and angle variables, $\{J'^M(\tilde{w}, Q^A), Q^A\}$, and a map to the worldline $Z[g_{\alpha\beta}] : \{J'^M, Q^A\} \to z^\mu$. The new action variables will be constructed to be independent of the periodic fast time variables Q^A. This construction will require a more complicated mapping $Z[g_{\alpha,\beta}](J'^M, Q^A)$ than the corresponding mapping for the original variables $\{J^M, q^A\}$, and in particular the function Z will now encode the remaining fast-time variation in the orbit.

Let us write out the explicit form of the action angle variable form of the multiscale system of equations:

$$\begin{aligned}\frac{dq_A}{dw} &= \omega_\alpha(J^M, \tilde{t}) + \varepsilon \bar{g}^{(1)}_A[\tilde{t}, P^M(\tilde{t})] + \varepsilon \tilde{g}^{(1)}_A[g_{\alpha\beta}(\tilde{t}, J^M, q^A)] + \mathcal{O}(\varepsilon^2)\\ \frac{dJ^M}{dw} &= \varepsilon \bar{G}^{M(1)}[\langle {}_s\psi(\tilde{t}, J^M)\rangle] + \varepsilon \tilde{G}^{M(1)}[g_{\alpha\beta}]^{(1)}(\tilde{t}, J^M, q^A)] + \mathcal{O}(\varepsilon^2)\end{aligned} \tag{5.84}$$

Following [350, 379], we suggest the 'near identity' transformation $\{P, q\} \to \{P', Q\}$,through $\mathcal{O}(\varepsilon^2)$:

$$\begin{aligned}J'^M &= J^M + \varepsilon T^M(J^M, Q^A, \tilde{t})\\ Q^A &= q^A + \varepsilon L^A(J^M, q^A, \tilde{t})\end{aligned} \tag{5.85}$$

We will separate these dynamics into a 'secular' piece and a 'oscillatory' piece, such that the secular force is given by the barred quantities $\bar{g}, \bar{G}$, and the oscillatory force is given by the tilded quantities $\tilde{g}, \tilde{G}$.

Flanagan and Vines [379] derive expressions for these transformations. In particular, they find the transformations

$$\begin{aligned} T^M_{K^A} =& \frac{i\tilde{G}^M_{K^A}}{K^A\Omega_A} \\ L^A_{K^A} =& \frac{i}{K^A\Omega_A}\left(\tilde{f} - \frac{\partial\omega^A}{\partial J_M}T_M\right), \end{aligned} \tag{5.86}$$

which will remove all dependence on the original oscillatory component of the forces, for any non-resonant orbit. This transforms the above equations to the form:

$$\begin{aligned} \frac{dQ_A}{dt} =& \omega_A(J^M,\tilde{t}) + \varepsilon\bar{g}^{(1)}_A[\tilde{t},P^M(\tilde{t})] + \mathcal{O}(\varepsilon^2) \\ \frac{dJ'^M}{dt} =& \varepsilon\bar{G}^{M(1)}[\langle {}_s\psi(\tilde{t},J^M)\rangle] + \mathcal{O}(\varepsilon^2). \end{aligned} \tag{5.87}$$

Consider, now, a more general set of gauge transformations for the action and action angle variables, in which the action angle variables are permitted to also have an $\mathcal{O}(1)$ slow-time dependent phase shift.

$$\begin{aligned} J'^M =& J^M + \varepsilon T^M(J^M,q^A,\tilde{t}) \\ Q'^A =& q^A + L'^A(\tilde{t}) + \varepsilon L^A(J^M,q^A,\tilde{t}) \end{aligned} \tag{5.88}$$

Note that any function which depended exclusively periodically on the original fast time variable at fixed $\tilde{t}$, will also depend exclusively periodically on the leading $q^A + L'^A(\tilde{t})$ at fixed $\tilde{t}$. Therefore, after the transformation, the equations of motion become

$$\begin{aligned} \frac{dQ'^A}{dt} =& \omega^A(J^M,\tilde{t}) - \varepsilon\partial_{\tilde{t}}L'^A + \varepsilon\bar{g}^{(1)A}[\tilde{t},P^M(\tilde{t})] + \mathcal{O}(\varepsilon^2), \\ \frac{dJ'^M}{dt} =& \varepsilon\bar{G}^{M(1)}[\langle {}_s\psi(\tilde{t},J^M)\rangle] + \mathcal{O}(\varepsilon^2). \end{aligned} \tag{5.89}$$

We may make the choice to take $L'^A = \varepsilon\bar{g}^{A(1)}[\tilde{t},P^M(\tilde{t})]$, in which case the equations of motion are reduced to

$$\begin{aligned} \frac{dQ'^A}{dt} =& \omega^A(J^M,\tilde{t}) + \mathcal{O}(\varepsilon^2), \\ \frac{dJ'^M}{dt} =& \varepsilon\bar{G}^{M(1)}[\langle {}_s\psi(\tilde{t},J^M)\rangle] + \mathcal{O}(\varepsilon^2). \end{aligned} \tag{5.90}$$

5.A.2 General treatment of inversion of $\{P'^M, Q'^A\}$ for the orbit

We start by discussing the general identities from which the equations of motion are derived. In each step, we will promote $P^M \to P^M(\tilde{w}, q^A)$. The goal is to have a set of generic formula for deriving the resulting worldline position, given the orbit parameters and action-angle variables.

For convenience, we will start with Boyer-Lindquist coordinates, and transform to the height-function based time,

$$w = t + h(r), \tag{5.91}$$

only where necessary.

Generically, we may write the defining equation for the Hamiltonian as

$$\mu^2 = g^{\mu\nu} p_\mu p_\nu, \tag{5.92}$$

and the velocities may be written

$$\frac{dz^\alpha}{d\tau} = \frac{1}{\mu} p_\beta g^{\alpha\beta}. \tag{5.93}$$

We now wish to re-write the components of z^α in terms of their derivatives with respect to w. We obtain

$$\frac{dz^w}{d\tau} = \frac{dz^t}{d\tau} + h'(r)\frac{dz^r}{d\tau}. \tag{5.94}$$

Therefore, the remaining important degrees of freedom for the worldline velocities may be re-written as

$$\frac{dz^i}{dw} = \frac{p_\beta g^{i\beta}}{p_\beta g^{t\beta} + h'(r) p_\beta g^{r\beta}}, \tag{5.95}$$

where i runs over r, θ, ϕ - the time coordinate derivative gives redundant information with the r coordinate once we have expressed the velocities as the first derivatives of z with respect to w.

We now perturbatively expand the worldline derivatives above to find,

$$\left(\frac{dz^i}{dw}\right)^{(1)} = \left(\frac{p_\beta g^{i\beta}}{p_\beta g^{t\beta} + h'(r) p_\beta g^{r\beta}}\right)^{(1)}, \tag{5.96}$$

To compute the function $z^{(1)i}$, we will require the quantities,

$$\begin{aligned}\Omega_A^{(0)}(P'^{(0)M}) &= \left\langle \left(\frac{\partial H}{\partial J^A}\right)^{(0)} \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}}\right)^{(0)} \right\rangle \\ &= \left(\frac{\partial H}{\partial J^A}\right)^{(0)} \left\langle \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}}\right)^{(0)} \right\rangle,\end{aligned} \tag{5.97a}$$

$$\Omega_A^{(1)}(P'^{(1)M}, P'^{(0)}) = \left\langle \left(\frac{\partial H}{\partial J^A}\right)^{(1)} \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}}\right)^{(0)} \right\rangle + \left(\frac{\partial H}{\partial J^A}\right)^{(0)} \left\langle \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}}\right)^{(1)} \right\rangle + \left\langle g_A(\tilde{w}, q^A, P^M) \right\rangle , \tag{5.97b}$$

$$\frac{\partial z^{(0)i}}{\partial P'^M} = \frac{\partial z^{(0)i}}{\partial P^M} + \mathcal{O}(\varepsilon) \tag{5.97c}$$

$$\left(\frac{p_\beta g^{i\beta}}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}}\right)^{(1)} = \left(\frac{-E g^{it} + L_z g^{i\phi} + p_\theta g^{i\theta} + p_r g^{ir}}{-E(g^{tt} + h'(r) g^{rt}) + L_z(g^{\phi t} + h'(r) g^{\phi r}) + p_\theta(g^{\theta t} + h'(r) g^{\theta r}) + p_r(g^{rt} + h'(r) g^{rr})}\right)^{(1)} . \tag{5.97d}$$

The frequency values $\partial H / \partial J^A$ are derived by inverting the Jacobian:

$$\begin{pmatrix} \frac{\partial J_t}{\partial H} & \frac{\partial J_t}{\partial E} & \frac{\partial J_t}{\partial L_z} & \frac{\partial J_t}{\partial Q} \\ \frac{\partial J_r}{\partial H} & \frac{\partial J_r}{\partial E} & \frac{\partial J_r}{\partial L_z} & \frac{\partial J_r}{\partial Q} \\ \frac{\partial J_\phi}{\partial H} & \frac{\partial J_\phi}{\partial E} & \frac{\partial J_\phi}{\partial L_z} & \frac{\partial J_\phi}{\partial Q} \\ \frac{\partial J_\theta}{\partial H} & \frac{\partial J_\theta}{\partial E} & \frac{\partial J_\theta}{\partial L_z} & \frac{\partial J_\theta}{\partial Q} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial H}{\partial J_t} & \frac{\partial H}{\partial J_r} & \frac{\partial H}{\partial J_\phi} & \frac{\partial H}{\partial J_\theta} \\ \frac{\partial E}{\partial J_t} & \frac{\partial E}{\partial J_r} & \frac{\partial E}{\partial J_\phi} & \frac{\partial E}{\partial J_\theta} \\ \frac{\partial L_z}{\partial J_t} & \frac{\partial L_z}{\partial J_r} & \frac{\partial L_z}{\partial J_\phi} & \frac{\partial L_z}{\partial J_\theta} \\ \frac{\partial Q}{\partial J_t} & \frac{\partial Q}{\partial J_r} & \frac{\partial Q}{\partial J_\phi} & \frac{\partial Q}{J_\theta} \end{pmatrix} = \mathbb{1}. \tag{5.98}$$

The desired quantities are enumerated in the subsequent sections for Schwarzschild and for Kerr.

5.A.3 | Kerr orbit quantities

In order to compute the remaining quantities, it will be useful to review the leading order map to worldine coordinates from geodesic parameters and angle variables. The leading order orbit equations of motion are [380]

$$p^t = \mu\Sigma \frac{dz^t}{d\tau} = \frac{r^2 + a^2}{\Delta}\left(E(r^2 + a^2) - aL_z\right) + a\left(aE\sin^2\theta - L_z\right) \tag{5.99a}$$

$$p^r = \mu\Sigma \frac{dz^r}{d\tau} = (\pm)^r \sqrt{V_r} \tag{5.99b}$$

$$p^\theta = \mu\Sigma \frac{dz^\theta}{d\tau} = (\pm)^\theta \sqrt{V_\theta} \tag{5.99c}$$

$$p^\phi = \mu\Sigma \frac{dz^\phi}{d\tau} = \frac{a}{\Delta}\left(E(r^2 + a^2) - aL_z\right) - aE + \frac{L_z}{\sin^2(\theta)}. \tag{5.99d}$$

The r and θ components have a sign ambiguity when written as first order differential equations, as the radial velocities and polar coordinate velocities will reach turning points triperiodically. For the purposes of these derivations, we will assume that the sign is consistently computed in q^A space, likely using continuity and zero-crossing properties of the potentials. We will therefore, for convenience, carry the (in general, distinct pair of) signs through the computation.

Due to the comparative ease of working with the equations of motion in proper time, we will first present the techniques for computing the motion as a function of the periodic variables using the frequencies with respect to proper time, which we define as

$$\frac{\partial q^A}{\partial \tau} = \Omega_\tau^A + \varepsilon g_\tau^{(1)A}(q^A, P^M) + \dots \tag{5.100}$$

The leading order functions to integrate are then

$$\mu\Sigma\Omega_\tau^A \frac{dz^t}{dq^A} = \frac{r^2+a^2}{\Delta}\left(E(r^2+a^2) - aL_z\right) + a\left(aE\sin^2\theta - L_z\right) \tag{5.101a}$$

$$\mu\Sigma\Omega_\tau^A \frac{dz^r}{dq^A} = (\pm)^r\sqrt{V_r} \tag{5.101b}$$

$$\mu\Sigma\Omega_\tau^A \frac{dz^\theta}{dq^A} = (\pm)^\theta\sqrt{V_\theta} \tag{5.101c}$$

$$\mu\Sigma\Omega_\tau^A \frac{dz^\phi}{dq^A} = \frac{a}{\Delta}\left(E(r^2+a^2) - aL_z\right) - aE + \frac{L_z}{\sin^2(\theta)}. \tag{5.101d}$$

The Kerr orbits form a troublesome coupled system of nonlinear, first-order differential equations. So, these must be integrated numerically. However, in principle this set of differential equations may be used to assemble a lookup table for $z^\mu(q^A, P^M)$, for a gridding of the fast-time space at each point in the space of geodesic orbits.

At subleading order, the equations of motion take the form

$$\Omega_\tau^A \frac{\partial z^{(1)\mu}}{\partial q^A} = \mathcal{M}^\mu_{\ \tau\nu}(q_\tau^A, P^{(0)M}) z^{(1)\nu} + \mathcal{N}^\mu_{\ \tau M}(q_\tau^A, P^{(0)M})\left(P'^{(1)M} - i\frac{G^M(P^{(0)}, q_\tau^A)}{k_A\Omega_\tau^A}\right) - \frac{\partial z^{(0)\mu}}{\partial q_\tau^A}\Omega_\tau^{(1)A}. \tag{5.102}$$

In principle, these functions $z^{(1)}$ can also be pre-computed for several points in the phase space $\{q_\tau^A, P^{(0)M}, P^{(1)M}\}$. It is possible that generating such cached results for the subleading worldline is prohibitively expensive, and such integration would need to be performed online.

The functions $\mathcal{M}_\tau$ and $\mathcal{N}_\tau$ are determined by taking derivatives of the leading equations of motion,

$$\mathcal{M}^t_{\tau\nu} = \partial_\nu \left(\frac{r^2+a^2}{\Delta} \left(E(r^2+a^2) - aL_z \right) + a \left(aE \sin^2\theta - L_z \right) \right) \tag{5.103a}$$

$$\mathcal{M}^r_{\tau\nu} = \partial_\nu \left((\pm)^r \sqrt{V_r} \right) \tag{5.103b}$$

$$\mathcal{M}^\theta_{\tau\nu} = \partial_\nu \left((\pm)^\theta \sqrt{V_\theta} \right) \tag{5.103c}$$

$$\mathcal{M}^\phi_{\tau\nu} = \partial_\nu \left(\frac{a}{\Delta} \left(E(r^2+a^2) - aL_z \right) - aE + \frac{L_z}{\sin^2(\theta)} \right), \tag{5.103d}$$

and

$$\mathcal{N}^t_{\tau M} = \frac{\partial}{\partial P^M} \left(\frac{r^2+a^2}{\Delta} \left(E(r^2+a^2) - aL_z \right) + a \left(aE \sin^2\theta - L_z \right) \right) \tag{5.104a}$$

$$\mathcal{N}^r_{\tau M} = \frac{\partial}{\partial P^M} \left((\pm)^r \sqrt{V_r} \right) \tag{5.104b}$$

$$\mathcal{N}^\theta_{\tau M} = \frac{\partial}{\partial P^M} \left((\pm)^\theta \sqrt{V_\theta} \right) \tag{5.104c}$$

$$\mathcal{N}^\phi_{\tau M} = \frac{\partial}{\partial P^M} \left(\frac{a}{\Delta} \left(E(r^2+a^2) - aL_z \right) - aE + \frac{L_z}{\sin^2(\theta)} \right). \tag{5.104d}$$

The Ω^A_τ are derived directly from the action-angle formalism,

$$\Omega^{(0)A}_\tau(P'^{(1)M}, P'^{(0)}) = \left\langle \left(\frac{\partial H}{\partial J^A} \right)^{(0)} \right\rangle \tag{5.105a}$$

$$\Omega^{(1)A}_\tau(P'^{(1)M}, P'^{(0)}) = \left\langle \left(\frac{\partial H}{\partial J^A} \right)^{(1)} \right\rangle + \left\langle g_{\tau A}(\tilde{w}, q^A, P^M) \right\rangle. \tag{5.105b}$$

For the purposes of use in the computation of the metric perturbations, frequencies with respect to the coordinate w are preferable. Those frequencies are related to the action-angle frequencies by multiplication and time-averaging associated with orbital values, which would be cheap to compute provided the worldline is already obtained via the above methods.

$$\Omega^{(1)}_A(P'^{(1)M}, P'^{(0)}) = \left(\frac{\partial H}{\partial J^A} \right)^{(0)} \left\langle \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}} \right)^{(0)} \right\rangle \tag{5.106a}$$

$$\begin{aligned} \Omega^{(1)}_A(P'^{(1)M}, P'^{(0)}) = &\left\langle \left(\frac{\partial H}{\partial J^A} \right)^{(1)} \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}} \right)^{(0)} \right\rangle \\ &+ \left(\frac{\partial H}{\partial J^A} \right)^{(0)} \left\langle \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}} \right)^{(1)} \right\rangle \\ &+ \left\langle g_{\tau A}(\tilde{w}, q^A, P^M) \left(\frac{1}{p_\beta g^{\beta t} + h'(r) p_\beta g^{\beta r}} \right)^{(0)} \right\rangle, \end{aligned} \tag{5.106b}$$

Consider now the functions $\Omega^{(0)M}(P^M)$ and $\Omega^{(1)M}(P^M)$. These frequencies are defined as the leading and first order contributions to the frequencies as functions of the geodesic parameters. The key distinction is that the same functional form holds, even if the geodesic parameters P^M are not constants, but instead functions of $\{q^A, \tilde{w}\}$.

The frequency values derived from the action variables $\partial H/\partial J^A$ are derived in [380], and reviewed in the appendix of [306]. The results require the definition of several functions,

$$W(P^M) = \int_{r_1}^{r_2} \frac{r^2 E(r^2+a^2) - 2Mra(L_z - aE)}{\Delta\sqrt{V_r}} dr \tag{5.107a}$$

$$X(P^M) = \int_{r_1}^{r_2} \frac{dr}{\sqrt{V_r}} \tag{5.107b}$$

$$Y(P^M) = \int_{r_1}^{r_2} \frac{r^2}{\sqrt{V_r}} dr \tag{5.107c}$$

$$Z(P^M) = \int_{r_1}^{r_2} \frac{r[L_z r - 2M(L_z - aE)]}{\Delta\sqrt{V_r}} dr \tag{5.107d}$$

, and elliptic integrals,

$$K(k) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2\sin^2(\theta)}}, \tag{5.108a}$$

$$E(k) = \int_0^{\pi/2} d\theta \sqrt{1 - k^2\sin^2(\theta)}, \tag{5.108b}$$

$$\Pi(\phi, n, k) = \int_0^{\phi} \frac{d\theta}{(1 - n\sin^2(\theta))\sqrt{1 - k^2\sin^2(\theta)}}, \tag{5.108c}$$

where $k = \sqrt{z_-/z_+}$ is the ratio of the two roots of $V_\theta(z) = 0$, for $z = \cos^2(\theta)$.

are that the freqencies take values,

$$\frac{\partial H}{\partial J_r} = \frac{\pi K(k)}{K(k)Y + a^2 z_+[K(k) - E(k)]X} \tag{5.109a}$$

$$\frac{\partial H}{\partial J_\theta} = \frac{\pi a\sqrt{z_+(\mu^2 - E^2)}X/2}{K(k)Y + a^2 z_+[K(k) - E(k)]X} \tag{5.109b}$$

$$\frac{\partial H}{\partial J_\phi} = \frac{K(k)Z + L_z[\Pi(\pi/2, z_-, k) - K(k)]X}{K(k)Y + a^2 z_+[K(k) - E(k)]X}. \tag{5.109c}$$

These give the desired frequencies as arbitrary functions of P^M. We may evaluate the leading order portion of these $\Omega^{(0)}(\tilde{w})$ by making the direct replacement $\{E, L_z, Q\} \to \{E^{(0)}(\tilde{w}), L_z^{(0)}(\tilde{w}), Q^{(0)}(\tilde{w})\}$.

The subleading corrections to the frequencies are straightforward to compute, but for completeness I quote the results in full,

$$\begin{aligned}\left(\frac{\partial H}{\partial J_r}\right)^{(1)} =& -\frac{\pi K^{(0)}(k)}{\left(K^{(0)}(k)Y^{(0)}+a^2 z_+^{(0)}[K^{(0)}(k)-E^{(0)}(k)]X^{(0)}\right)^2}\\ &\times\Big(K^{(0)}(k)Y^{(1)}+K^{(1)}(k)Y^{(0)}+a^2 z_+^{(1)}[K^{(0)}(k)-E^{(0)}(k)]X^{(0)}\\ &\quad+a^2 z_+^{(0)}[K^{(0)}(k)-E^{(0)}(k)]X^{(1)}+a^2 z_+^{(0)}[K^{(1)}(k)-E^{(1)}(k)]X^{(0)}\Big)\\ &+\frac{\pi K^{(1)}(k)}{K^{(0)}(k)Y^{(0)}+a^2 z_+^{(0)}[K^{(0)}(k)-E^{(0)}(k)]X^{(0)}}\end{aligned} \tag{5.110a}$$

$$\begin{aligned}\left(\frac{\partial H}{\partial J_\theta}\right)^{(1)} =& \frac{1}{2K^{(0)}(k)}a\sqrt{z_+^{(0)}(\mu^2-E^{(0)2})}\left(X^{(1)}\left(\frac{\partial H}{\partial J_r}\right)^{(0)}+X^{(0)}\left(\frac{\partial H}{\partial J_r}\right)^{(1)}\right)\\ &+\frac{a(\mu^2-E^{(0)2})z_+^{(1)}+2az_+^{(0)}E^{(0)}E^{(1)}}{4K^{(0)}(k)\sqrt{z_+^{(0)}(\mu^2-E^{(0)2})}}X^{(0)}\left(\frac{\partial H}{\partial J_r}\right)^{(0)}\\ &-\frac{K^{(1)}(k)}{2K^{(0)2}(k)}a\sqrt{z_+^{(0)}(\mu^2-E^{(0)2})}X^{(0)}\left(\frac{\partial H}{\partial J_r}\right)^{(0)}\end{aligned} \tag{5.110b}$$

$$\begin{aligned}\left(\frac{\partial H}{\partial J_\phi}\right)^{(1)} =& \left(\frac{\partial H}{\partial J_r}\right)^{(0)}\Big(Z^{(1)}/\pi+L_z^{(1)}\frac{\Pi^{(0)}(\pi/2,z_-,k)-K^{(0)}(k)}{\pi K^{(0)}(k)}X^{(0)}\\ &\quad+L_z^{(0)}\frac{\Pi^{(0)}(\pi/2,z_-,k)-K^{(0)}(k)}{\pi K^{(0)}(k)}X^{(1)}\\ &\quad+L_z^{(0)}\frac{\Pi^{(1)}(\pi/2,z_-,k)-K^{(1)}(k)}{\pi K(k)}X^{(0)}\\ &\quad-L_z^{(0)}K^{(1)}(k)\frac{\Pi^{(0)}(\pi/2,z_-,k)-K^{(0)}(k)}{\pi(K^{(0)}(k))^2}X^{(0)}\Big)\\ &+\left(\frac{\partial H}{\partial J_r}\right)^{(1)}\left(Z^{(0)}/\pi+L_z^{(0)}\frac{\Pi(\pi/2,z_-,k)-K(k)}{\pi K(k)}X^{(0)}\right).\end{aligned} \tag{5.110c}$$

We also need the perturbed auxillary functions,

$$
\begin{aligned}
&W^{(1)}(P^M)\\
&\quad=\int_{r_1}^{r_2}\left(\frac{r^2E^{(1)}(r^2+a^2)-2Mra(L_z^{(1)}-aE^{(1)})}{\Delta\sqrt{V_r^{(0)}}}-\frac{r^2E^{(0)}(r^2+a^2)-2Mra(L_z^{(0)}-aE^{(0)})}{2\Delta\left(V_r^{(0)}\right)^{3/2}}\right)dr\\
&\qquad+\frac{r_2^{(1)}\left(r_2^{(0)2}E^{(1)}(r_2^{(0)2}+a^2)-2Mr_2^{(0)}a(L_z^{(1)}-aE^{(1)})\right)}{\Delta(r_2^{(0)})\sqrt{V_r^{(0)}(r_2^{(0)})}}\\
&\qquad-\frac{r_1^{(1)}\left(r_1^{(0)2}E^{(1)}(r_1^{(0)2}+a^2)-2Mr_1^{(0)}a(L_z^{(1)}-aE^{(1)})\right)}{\Delta(r_1^{(0)})\sqrt{V_r^{(0)}(r_1^{(0)})}}
\end{aligned}
\tag{5.111a}
$$

$$
X^{(1)}(P^M)=\int_{r_1}^{r_2}-\frac{V_r^{(1)}dr}{\left(V_r^{(0)}\right)^{3/2}}+\frac{r_2^{(1)}}{\sqrt{V_r^{(0)}(r_2^{(0)})}}-\frac{r_1^{(1)}}{\sqrt{V_r^{(0)}(r_1^{(0)})}}
\tag{5.111b}
$$

$$
Y^{(1)}(P^M)=\int_{r_1}^{r_2}-\frac{r^2V_r^{(1)}}{2\left(V_r^{(0)}\right)^{3/2}}dr+\frac{r_2^{(1)}r_2^{(0)2}}{\sqrt{V_r^{(0)}(r_2^{(0)})}}-\frac{r_1^{(1)}r_1^{(0)2}}{\sqrt{V_r^{(0)}(r_1^{(0)})}}
\tag{5.111c}
$$

$$
\begin{aligned}
Z^{(1)}(P^M)=&\int_{r_1}^{r_2}\frac{r[L_z^{(1)}r-2M(L_z^{(1)}-aE^{(1)})]}{\Delta\sqrt{V_r^{(0)}}}-\frac{r[L_z^{(0)}r-2M(L_z^{(0)}-aE^{(0)})]V_r^{(1)}}{2\Delta\left(V_r^{(0)}\right)^{3/2}}dr\\
&+\frac{r_2^{(1)}r_2^{(0)}[L_z^{(0)}r_2^{(0)}-2M(L_z^{(0)}-aE^{(0)})]}{\Delta(r_2^{(0)})\sqrt{V_r^{(0)}(r_2^{(0)})}}-\frac{r_1^{(1)}r_1^{(0)}[L_z^{(0)}r_1^{(0)}-2M(L_z^{(0)}-aE^{(0)})]}{\Delta(r_1^{(0)})\sqrt{V_r^{(0)}(r_1^{(0)})}},
\end{aligned}
\tag{5.111d}
$$

and the perturbed elliptic integrals,

$$
K^{(1)}(k)=\int_0^{\pi/2}-\frac{k^{(0)}k^{(1)}d\theta}{(1-k^{(0)2}\sin^2\theta)^{3/2}}
\tag{5.112a}
$$

$$
E^{(1)}(k)=\int_0^{\pi/2}\frac{k^{(0)}k^{(1)}d\theta}{\sqrt{1-k^{(0)2}\sin^2\theta}}
\tag{5.112b}
$$

$$
\begin{aligned}
\Pi^{(1)}(\pi/2,z_-,k)=\int_0^{\pi/2}\Bigg(&-\frac{z_-^{(1)}\sin^2\theta}{(1-z_-^{(0)}\sin^2\theta)^2\sqrt{1-k^{(0)2}\sin^2\theta}}\\
&-\frac{k^{(0)}k^{(1)}}{(1-z_-^{(0)}\sin^2\theta)(1-k^{(0)2}\sin^2\theta)^{3/2}}\Bigg)d\theta
\end{aligned}
\tag{5.112c}
$$

$$
k^{(1)}=\frac{z_-^{(1)}}{2z_-^{(0)}}k^{(0)}-\frac{z_+^{(1)}}{2z_+^{(0)}}k^{(0)}
\tag{5.112d}
$$

The potentials and extremal coordinates that appear are then

$$\begin{aligned}
V_r^{(1)}(r) =& 2[(r^2+a^2)E^{(0)} - aL_z^{(0)}][(r^2+a^2)E^{(1)} - aL_z^{(1)}] \\
&- \Delta[\mu^2 r^2 + 2(L_z^{(0)} - aE^{(0)})(L_z^{(1)} - aE^{(1)}) + Q^{(1)}] && \text{(5.113a)} \\
r_2^{(1)} =& V_r^{(1)}(r_2^{(0)}) \left(\partial_r V_r^{(0)}(r_2^{(0)})\right)^{-1} \\
=& V_r^{(1)}(r_2^{(0)}) \Big\{ 4r_2^{(0)}[(r_2^{(0)2}+a^2)E^{(0)} - aL_z^{(0)}] - 2\mu^2 r_2^{(0)}\Delta(r_2^{(0)}) \\
&- 2(r_2^{(0)} - M)\left(\mu^2 r_2^{(0)2} + (L_z^{(0)} - aE^{(0)})^2 + Q^{(0)}\right) \Big\} && \text{(5.113b)} \\
r_1^{(1)} =& V_r^{(1)}(r_1^{(0)}) \left(\partial_r V_r^{(0)}(r_1^{(0)})\right)^{-1} \\
=& V_r^{(1)}(r_1^{(0)}) \Big\{ 4r_1^{(0)}[(r_1^{(0)2}+a^2)E^{(0)} - aL_z^{(0)}] - 2\mu^2 r_1^{(0)}\Delta(r_1^{(0)}) \\
&- 2(r_1^{(0)} - M)\left(\mu^2 r_1^{(0)2} + (L_z^{(0)} - aE^{(0)})^2 + Q^{(0)}\right) \Big\} && \text{(5.113c)} \\
V_\theta^{(1)}(z) =& Q^{(1)} - \left(-2E^{(0)}E^{(1)}a^2 + \frac{2L_z^{(0)}L_z^{(1)}}{1-z}\right) z && \text{(5.113d)} \\
z_+^{(1)} =& V_\theta^{(1)}(z_+^{(0)}) \left(\partial_z V_\theta^{(0)}(z_+^{(0)})\right)^{-1} = \frac{V_\theta^{(1)}(z_+^{(0)})}{(\mu^2 - E^{(0)2})a^2 + \frac{L_z^{(0)2}}{1-z_+^{(0)}} + \frac{L_z^{(0)2} z_+^{(0)}}{(1-z_+^{(0)})^2}} && \text{(5.113e)} \\
z_-^{(1)} =& V_\theta^{(1)}(z_-^{(0)}) \left(\partial_z V_\theta^{(0)}(z_-^{(0)})\right)^{-1} = \frac{V_\theta^{(1)}(z_-^{(0)})}{(\mu^2 - E^{(0)2})a^2 + \frac{L_z^{(0)2}}{1-z_-^{(0)}} + \frac{L_z^{(0)2} z_-^{(0)}}{(1-z_-^{(0)})^2}} && \text{(5.113f)}
\end{aligned}$$

5.A.4 | Transformation from J^M to P^M

It is more convenient to work exclusively with the geodesic parameters E, L_z, Q in terms of the fast and slow time quantities. The above equations of motion can just as easily be treated using these as the evolving parameters, from which we will derive the orbit. During the computation, we assume that a corresponding set of functions exist. Here we give explicitly the functions required to move from the action variables to the geodesic parameters for generic orbits in Kerr.

If we have a derived value for the forcing functions for $\partial J^M/\partial w$, we can obtain the forcing functions $\partial P^M/\partial w$ via

$$\frac{\partial P^M}{\partial w} = \frac{\partial P^M}{\partial J^N}\frac{\partial J^N}{\partial w}, \tag{5.114}$$

where $\partial P^M/\partial J^N$ can be determined by inverting

$$\frac{\partial J^M}{\partial P^N}\frac{\partial P^N}{\partial J^K} = \delta^M{}_K \tag{5.115}$$

These quantities have been computed in generality by [Schmidt], giving the resulting derivatives,

$$\frac{\partial E}{\partial J^t} = -1, \qquad \frac{\partial L_z}{\partial J^t} = 0, \qquad \frac{\partial Q}{\partial J_t} = \frac{-2a^2 z_+^2 \left(K(k)W(P^M)(K(k)-E(k))\right)}{K(k)\left(K(k)Y(P^M)+a^2 z_+^2(K(k)-E(k))X(P^M)\right)} \tag{5.116a}$$

$$\frac{\partial E}{\partial J^r} = 0, \qquad \frac{\partial L_z}{\partial J^r} = 0, \qquad \frac{\partial Q}{\partial J^r} = \frac{2a^2 z_+^2 \pi(K(k)-E(k))}{K(k)Y(P^M)+a^2 z_+^2(K(k)-E(k))X(P^M)} \tag{5.116b}$$

$$\frac{\partial E}{\partial J^\theta} = 0, \qquad \frac{\partial L_z}{\partial J^\theta} = 0,$$
$$\frac{\partial Q}{\partial J^\theta} = \frac{-2\pi a^2 \beta z_+^3 (K(k)-E(k))X(P^M)}{2K(k)\left(K(k)Y(P^M)+a^2 z_+^2(K(k)-E(k))X(P^M)\right)} + \frac{\pi}{2}\frac{\beta z_+}{K(k)} \tag{5.116c}$$

$$\frac{\partial E}{\partial J^\phi} = 1, \qquad \frac{\partial L_z}{\partial J^\phi} = 0,$$
$$\frac{\partial Q}{\partial J^\phi} = \frac{2z_+^2\left(L_z(\Pi(\pi/2,z_-,k)-K(k))Y(P^M)-a^2(K(k)-E(k))Z(P^M)\right)}{K(k)Y(P^M)+a^2 z_+^2(K(k)-E(k))X(P^M)} \tag{5.116d}$$

We also note that the only one of these maps that is at all complicated is that associated with the Carter constant. Overall, it is easy to notice that limiting consideration to equatorial orbits will vastly simplify the reasoning associated with the orbit evolution.

5.B | Supplementary material for the flux identities in multiscale

5.B.1 | First order in multiscale formalism

First, I'd like to include a brief note on the choice of time coordinate. In the multiscale approximation that we'll present for use with the Einstein field equation, we will wish to evaluate the orbit-averaged derivative with respect to a convenient coordinate time

$$\left\langle \frac{d\mathcal{E}}{dw} \right\rangle = \left\langle \frac{d\tau_0}{dw}\frac{d\mathcal{E}}{d\tau_0} \right\rangle \tag{5.117}$$

For the purposes of this derivation, I will neglect the leading $d\tau/dw$ term. This should not make a substantial difference for the derivation, whenever I integrate by parts to obtain a total derivative with respect to τ, the leading correction will convert this to a total derivative with respect to w, which vanishes.

At first order, we have

$$\partial_{\tilde{\tau}_0}\mathcal{E}^{(0)} + \frac{D}{d\tau_0}\mathcal{E}^{(1)} = \xi_\alpha \partial_{\tilde{\tau}_0} u^{(0)\alpha} + \xi_\alpha \frac{D^2}{d\tau_0^2} z_{1\perp}^\alpha + \xi_\alpha \frac{D}{d\tau_0}\partial_{\tilde{\tau}_0} z_0^\alpha + u_0^\alpha \frac{D z_{1\perp}^\beta}{d\tau_0}\nabla_\alpha \xi_\beta + u_0^\alpha \partial_{\tilde{\tau}_0} z_0^\beta \nabla_\alpha \xi_\beta. \quad (5.118)$$

The right-hand side is similar to what we should expect from the first order self acceleration, expanded according to the multiscale variables.

The acceleration, given instead in terms of multiscale values reads:

$$\frac{D^2}{d\tau_0^2} z_{1\perp}^\mu + \partial_{\tilde{\tau}_0} u_0^\mu + \frac{D}{d\tau_0}\partial_{\tilde{\tau}_0} z_0^\mu = F^\mu + R^\mu{}_{\alpha\nu\beta} u^\alpha z_{1\perp}^\nu u^\beta \quad (5.119)$$

Performing a substitution with the energy equation, we obtain:

$$\langle \partial_{\tilde{w}} \mathcal{E}^{(0)} \rangle = \left\langle \frac{\partial \tau_0}{\partial w} \xi_\mu F_1^\mu \right\rangle + \left\langle \frac{\partial \tau_0}{\partial w} u_0^\alpha \partial_{\tilde{\tau}_0} z_0^\beta \nabla_\alpha \xi_\beta \right\rangle. \quad (5.120)$$

We can use the fact that the slow time derivative of the worldline will depend exclusively on the conserved quantities to re-write the second term on the right hand side:

$$\langle \partial_{\tilde{w}} \mathcal{E}^{(0)} \rangle - \partial_{\tilde{\tau}_0} P_0^M \left\langle \frac{\partial \tau_0}{\partial w} u_0^\alpha \frac{\partial z_0^\beta}{\partial P_0^M} \nabla_\alpha \xi_\beta \right\rangle = \left\langle \frac{\partial \tau_0}{\partial w} \xi_\mu F_1^\mu \right\rangle. \quad (5.121)$$

The second term on the right-hand side depends only on the slow time derivative of the worldline parameters and their instantaneous values, not on the metric perturbation. This introduces the mild inconvenience that the flux formulas are now a linear system of equations for the first derivatives of the conserved quantities, rather than direct linear equations.

The remaining term on the right-hand side can be simplified via the same steps used to simplify the expression in the first order self consistent computation, so the final result for the multiscale approximation at first order is

$$\left\langle \partial_{\tilde{w}} \mathcal{E}^{(0)} \right\rangle - \partial_{\tilde{\tau}_0} P_0^M \left\langle \frac{\partial \tau_0}{\partial w} u_0^\alpha \frac{\partial z_0^\beta}{\partial P_0^M} \nabla_\alpha \xi_\beta \right\rangle = \left\langle \frac{\partial \tau_0}{\partial w} u^\alpha u^\beta \partial_\xi h_{\alpha\beta}^{(1)} \right\rangle \quad (5.122)$$

Finally, the term on the right-hand side can once again be expressed in terms of the squares of the homogeneous amplitudes, as the $\partial\tau_0/\partial w$ should cancel with the integral measure to give the correct integral over the worldline, giving a right-hand side identical to (5.61)

5.B.2 Second order in the multiscale formalism

The second-order expansion proceeds by similar steps, albeit with more complicated expressions

$$\partial_{\tilde{\tau}_0}\mathcal{E}^{(1)}+\frac{D}{d\tau_0}\mathcal{E}^{(2)}=\xi_\alpha\partial_{\tilde{\tau}_0}\frac{D}{d\tau_0}z^\alpha_{1\perp}+\xi_\alpha\partial^2_{\tilde{\tau}_0}z^\alpha_0+\xi_\alpha\frac{D^2}{d\tau_0}z^\alpha_{2\ddagger}+\xi_\alpha\frac{D}{d\tau_0}\partial_{\tilde{\tau}_0}z^\alpha_{1\perp}+u^\alpha_0\frac{D}{d\tau_0}z^\beta_{2\ddagger}\nabla_\alpha\xi_\beta+u^\alpha_0\partial_{\tilde{\tau}_0}z^\beta_{1\perp}\nabla_\alpha\xi_\beta \tag{5.123}$$

The right-hand side is again equivalent to the self-acceleration at second order. Taking the expressions directly from [381], we have

$$\begin{aligned}\partial_{\tilde{\tau}_0}\mathcal{E}^{(1)}+\frac{D}{d\tau_0}\mathcal{E}^{(2)}=&\xi_\alpha F^\alpha_2-\xi_\mu P^\mu_0{}_\rho R^\rho{}_{\alpha\nu\beta}\partial_{\tilde{\tau}_0}z^\alpha_0u^\beta_0z^\nu_{1\perp}-\xi_\mu R^\mu{}_{\alpha\nu\beta}u^\alpha_0z^\nu_{1\perp}\partial_{\tilde{\tau}_0}z^\beta_0+u^\alpha_0\partial_{\tilde{\tau}_0}z^\beta_{1\perp}\nabla_\alpha\xi_\beta\\&-2\xi_\alpha P^\alpha_0{}_\rho R^\rho{}_{\mu\beta\nu}\frac{Dz^\mu_{1\perp}}{d\tau_0}z^\beta_{1\perp}u^\nu_0+2\xi_\alpha P^\alpha_0{}_\rho R^\rho{}_{\mu\beta\nu;\gamma}z^{(\mu}_{1\perp}u^{\beta)}_0z^{[\nu}_{1\perp}u^{\gamma]}_0\end{aligned} \tag{5.124}$$

Most of these terms can be evaluated given only instantaneous in $\tilde{\tau}_0$ knowledge of the first order worldline. However, to perform the trick we used for the first order self force which permitted simplification, we need to assume we've performed the near-identity transformation to ensure that the geodesic parameters depend exclusively on slow time.

The near-identity transformation comes at the cost of rolling first order metric dependence into the map $z^{(1)}[P^M,q^A]$. This will further complicate the final expression for the first-order metric contributions to the slowly evolving $\mathcal{E}^{(1)}$, but shouldn't change the central thesis that the second order perturbations contribute only in a way expressible as integrals over first order quantities and asymptotic fluxes.

Therefore, the time average of the fourth term on the right hand side can be re-expressed as

$$\left\langle u^\alpha\partial_{\tilde{\tau}_0}z^\beta_{1\perp}\nabla_\alpha\xi_\beta\right\rangle=\frac{\partial P^M_0}{\partial\tilde{\tau}_0}\left\langle u^\alpha\frac{\partial z^\beta_{1\perp}}{\partial P_0}\nabla_\alpha\xi_\beta\right\rangle+\frac{\partial P^M_1}{\partial\tilde{\tau}_0}\left\langle u^\alpha\frac{\partial z^\beta_{1\perp}}{\partial P_1}\nabla_\alpha\xi_\beta\right\rangle \tag{5.125}$$

The remaining terms on the right-hand side of (5.124) can be simplified a bit by integration by parts. The resulting equation is

$$\begin{aligned}\left\langle\partial_{\tilde{\tau}_0}\mathcal{E}^{(1)}\right\rangle=&\langle\xi_\alpha F^\alpha_2\rangle+\frac{\partial P^M_0}{\partial\tilde{\tau}_0}\left\langle u^\alpha\frac{\partial z^\beta_{1\perp}}{\partial P_0}\nabla_\alpha\xi_\beta\right\rangle+\frac{\partial P^M_1}{\partial\tilde{\tau}_0}\left\langle u^\alpha\frac{\partial z^\beta_{1\perp}}{\partial P_1}\nabla_\alpha\xi_\beta\right\rangle\\&-\left\langle\mathcal{E}^{(0)}u^\rho_0R_{\rho\alpha\nu\beta}u^\alpha_1z^\nu_{1\perp}u^\beta_0\right\rangle+2\left\langle\xi_\rho R^\rho{}_{\alpha\beta\nu}\partial_{\tilde{\tau}_0}z^{[\alpha}_0u^{\beta]}_0z^\nu_{1\perp}\right\rangle+\left\langle\xi_\rho R^\rho{}_{[\mu\beta]\nu}z^\mu_{1\perp}\frac{Dz^\beta_{1\perp}}{d\tau_0}u^\nu_0\right\rangle\end{aligned} \tag{5.126}$$

Focusing now on the expansion of the force term $\langle \xi_\alpha F_2^\alpha \rangle$:

$$\begin{aligned}\xi_\alpha F_2^\alpha =& -\frac{1}{2}\xi_\mu P_0^{\mu\rho}(2h^{(2)}_{\rho\sigma;\lambda} - h^{(2)}_{\sigma\lambda;\rho})u_0^\sigma u_0^\lambda - \frac{1}{2}\xi_\mu P_0^{\mu\rho}(2h^{(1)}_{\rho\sigma;\lambda\delta} - h^{(1)}_{\sigma\lambda;\rho\delta})u_0^\sigma u_0^\lambda z_{1\lambda}^\delta \\ &- \frac{1}{2}h^{(1)}_{\nu\sigma;\lambda}\xi_\mu \frac{Dz_{1\perp}^\mu}{d\tau_0}u_0^\nu u_0^\sigma u_0^\lambda - \xi_\mu h^{(1)}_{\nu\sigma;\lambda}P_0^{\mu\nu}\frac{Dz_{1\perp}^\sigma}{d\tau_0}u_0^\lambda - \xi_\mu P_0^{\mu\nu}h^{(1)}_{\nu\sigma;\lambda}u_0^\sigma \frac{Dz_{1\perp}^\lambda}{d\tau_0} \\ &+ \xi_\mu P_0^{\mu\nu}h^{(1)}_{\sigma\lambda;\nu}\frac{Dz_{1\perp}^\sigma}{d\tau_0}u_0^\lambda + \frac{1}{2}\xi_\mu P_0^{\mu\nu}h_\nu^{(1)\rho}(2h^{(1)}_{\rho\sigma;\lambda} - h^{(1)}_{\sigma\lambda;\rho})u_0^\sigma u_0^\lambda - \frac{1}{2}\xi_\mu \partial_{\tilde\tau} z_0^\mu u_0^\delta (h_{\delta\beta;\gamma}u_0^\beta u_0^\gamma) \\ &- \frac{1}{2}\xi_\mu u_0^\mu \partial_{\tilde\tau} z_0^\delta (2h^{(1)}_{\delta\beta;\gamma} - h^{(1)}_{\beta\gamma;\delta})u^\beta \partial_{\tilde\tau} z_0^\gamma - \xi_\mu P_0^{\mu\delta}\partial_{\tilde\tau}h^{(1)}_{\delta\beta}u^\beta \end{aligned} \tag{5.127}$$

By playing various tricks with integrating by parts and eliminating total time derivatives, this can be moderately condensed. The angle-averaged force expression simplifies to:

$$\begin{aligned}\langle \xi_\alpha F_2^\alpha \rangle =& \frac{1}{2}\left\langle u_0^\nu u_0^\nu \partial_\xi h^{(2)}_{\mu\nu}\right\rangle + \frac{1}{8}\left\langle u^\mu u^\nu u^\alpha u^\beta \partial_\xi (h^{(1)}_{\mu\nu}h^{(1)}_{\alpha\beta})\right\rangle + \left\langle \partial_\xi h^{(1)}_{\beta\gamma}u_1^\beta u_0^\gamma\right\rangle \\ &- \partial_{\tilde\tau_0}\left\langle u^\beta(\xi^\alpha - \frac{1}{2}\mathcal{E}^{(0)}u^\alpha)h^{(1)}_{\beta\alpha}\right\rangle - \frac{1}{2}\left\langle \mathcal{E}^{(0)}u^\alpha u^\beta \partial_{\tilde\tau_0}h_{\alpha\beta}\right\rangle - \left\langle \mathcal{E}^{(0)}h^{(1)}_{\sigma\lambda}\frac{D^2 z_{1\perp}^\sigma}{d\tau_0^2}u_0^\lambda\right\rangle \\ &- \frac{1}{2}\left\langle \mathcal{E}^{(0)}u_0^\delta u_0^\beta u_1^\gamma h_{\delta\beta;\gamma}\right\rangle - \left\langle \xi^\delta h^{(1)}_{\delta\beta;\gamma}u^\beta \partial_{\tilde\tau_0}z_0^\gamma\right\rangle + \left\langle \mathcal{E}^{(0)}u^\rho h^{(1)}_{\nu\sigma}R^\nu{}_{\rho\lambda\delta}u_0^\lambda u_0^\sigma z_{1\perp}^\delta\right\rangle \\ &+ 2\left\langle \xi^\rho h^{(1)}_{\nu\sigma}R^\nu{}_{(\rho\lambda)\delta}u_0^\sigma u_0^\lambda z_{1\perp}^\delta\right\rangle + \left\langle \xi_\mu R^\mu{}_{\alpha\nu\beta}u_0^\alpha z_{1\perp}^\nu u_0^\beta u_0^\gamma u_0^\delta h^{(1)}_{\gamma\delta}\right\rangle \\ &- \frac{1}{2}\left\langle \xi^\gamma{}_{;\delta}h^{(1)}_{\sigma\lambda;\gamma}u_0^\sigma u_0^\lambda \frac{Dz_{1\perp}^\delta}{d\tau_0}\right\rangle - \left\langle \xi^\sigma{}_{;\beta}h^{(1)}_{\sigma\gamma}u_1^\beta u_0^\gamma\right\rangle, \end{aligned} \tag{5.128}$$

note that the u_1 here refers to the full multiscale first order velocity $\partial_{\tilde\tau_0}z_0 + D_{\tau_0}z_{1\perp}$.

This is then combined with the remaining Riemann terms at this order (5.126) to find the the final expression for the loss of energy in terms of second order fluxes and the various first order terms, depending on the perturbed worldline and first order quantities.

The additional argument required to express $\langle u^\mu u^\nu \partial_\xi h^{(2)}_{\mu\nu}\rangle$ in terms of fluxes and a first order integral follows the same steps as the second order computation for the Self Consistent formalism, with the time components of the integration replaced by fast-time integration.

5.B.3 | Gauge transformation of the second-order flux expressions

We'd like to express the gauge characteristics of the quantity given in the previous subsection as the total flux

$$\begin{aligned}\left\langle \frac{d\mathcal{E}}{d\tau}\right\rangle =& \varepsilon\frac{1}{2}\left\langle u^\alpha u^\beta \mathcal{L}_\xi h^{R(1)}_{\alpha\beta}\right\rangle + \varepsilon^2\frac{1}{2}\left\langle u^\alpha u^\beta \mathcal{L}_\xi h^{R(2)}_{\alpha\beta}\right\rangle + \frac{1}{4}\varepsilon^2\left\langle u^\alpha u^\beta u^\gamma u^\delta \mathcal{L}_\xi h^{R(1)}_{\alpha\beta}h^{R(1)}_{\gamma\delta}\right\rangle \\ &- \varepsilon^2\partial_{\tilde\tau}\left\langle \xi^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma}\right\rangle - \frac{1}{2}\varepsilon^2\partial_{\tilde\tau}\left\langle \mathcal{E}u^\beta u^\gamma h^{(1)\mathcal{R}}_{\beta\gamma}\right\rangle + \mathcal{O}(\varepsilon^3) \end{aligned} \tag{5.129}$$

We now re-write the orbit averages in terms of the induced parameter s, rather than in terms of τ, so that we can take better advantage of the gauge transformation identities of $\dot{z}^\alpha$.

$$\begin{aligned}\left\langle \frac{d\mathcal{E}}{d\tau} \right\rangle =& \varepsilon \frac{1}{2} \left\langle \left(\frac{ds}{d\tau} \right) \dot{z}^\alpha \dot{z}^\beta \mathcal{L}_\xi h^{R(1)}_{\alpha\beta} \right\rangle_s + \varepsilon^2 \frac{1}{2} \left\langle \left(\frac{ds}{d\tau} \right) \dot{z}^\alpha \dot{z}^\beta \mathcal{L}_\xi h^{R(2)}_{\alpha\beta} \right\rangle_s \\ &+ \frac{1}{4} \varepsilon^2 \left\langle \left(\frac{ds}{d\tau} \right)^3 \dot{z}^\alpha \dot{z}^\beta \dot{z}^\gamma u^\delta \mathcal{L}_\xi h^{R(1)}_{\alpha\beta} h^{R(1)}_{\gamma\delta} \right\rangle_s - \partial_{\bar{\tau}} \left\langle \xi^\beta \dot{z}^\gamma h^{R(1)}_{\beta\gamma} \right\rangle_s \\ &- \frac{1}{2} \partial_{\bar{\tau}} \left\langle \left(\frac{ds}{d\tau} \right) \mathcal{E} \dot{z}^\beta \dot{z}^\gamma h^{R(1)}_{\beta\gamma} \right\rangle_s + \mathcal{O}(\varepsilon^3), \end{aligned} \tag{5.130}$$

and for reference note that

$$\mathcal{L}_\zeta \left(\frac{ds}{d\tau} \right) = \left(\frac{ds}{d\tau} \right)^3 \dot{z}^\alpha \dot{z}^\beta \nabla_\alpha \zeta_\beta = \frac{1}{2} \left(\frac{ds}{d\tau} \right)^3 \dot{z}^\alpha \dot{z}^\beta \mathcal{L}_\zeta g^{(0)}_{\alpha\beta} \tag{5.131}$$

To perform this computation more easily, note also the identity

$$\Delta(AB) = B\Delta A + A\Delta B + \Delta A \Delta B \tag{5.132}$$

So, it will be advantageous to compute directly the gauge change of $u^\mu u^\nu \mathcal{L}_\xi h^R_{\mu\nu}$ at first and second orders.

$$\begin{aligned}\Delta \left(\dot{z}^\mu \dot{z}^\nu \left(\frac{ds}{d\tau} \right) \mathcal{L}_\xi h^{R(1)}_{\mu\nu} \right) =& \left(\frac{ds}{d\tau} \right) \dot{z}^\mu \dot{z}^\nu \mathcal{L}_\xi \mathcal{L}_\zeta g^{(0)}_{\mu\nu} - \varepsilon \dot{z}^\mu \dot{z}^\nu \left(\frac{ds}{d\tau} \right) (\mathcal{L}_\zeta + \pounds_\zeta) \left(\mathcal{L}_\xi (h^{R(1)}_{\mu\nu}(z(\tau); z)) \right) \\ &- \varepsilon \dot{z}^\mu \dot{z}^\nu \left(\frac{ds}{d\tau} \right) \mathcal{L}_\zeta \left(\mathcal{L}_\xi \mathcal{L}_\zeta g^{(0)}_{\mu\nu} \right) \\ &- \frac{1}{2} \varepsilon \left(\frac{ds}{d\tau} \right)^3 \dot{z}^\alpha \dot{z}^\beta \mathcal{L}_\zeta g^{(0)}_{\alpha\beta} \dot{z}^\mu \dot{z}^\nu \left(\mathcal{L}_\xi h_{\mu\nu} + \mathcal{L}_\xi \mathcal{L}_\zeta g^{(0)}_{\mu\nu} \right) + \mathcal{O}(\varepsilon^2) \end{aligned} \tag{5.133a}$$

$$\Delta \left(\left(\frac{ds}{d\tau} \right) \dot{z}^\mu \dot{z}^\nu \mathcal{L}_\xi (h^{R(2)}_{\mu\nu}(z(\tau); z) \right) = \left(\frac{ds}{d\tau} \right) \dot{z}^\mu \dot{z}^\nu \mathcal{L}_\xi \left(\frac{1}{2} \mathcal{L}^2_\zeta g_{\mu\nu} + (\mathcal{L}_\zeta + \pounds_\zeta) h^{R(1)}_{\mu\nu} \right) + \mathcal{O}(\varepsilon) \tag{5.133b}$$

$$\Delta \left(\left(\frac{ds}{d\tau} \right) \dot{z}^\mu \dot{z}^\nu h^{R(1)}_{\mu\nu} \right) = \left(\frac{ds}{d\tau} \right) \dot{z}^\mu \dot{z}^\nu \mathcal{L}_\zeta g^{(0)}_{\mu\nu} + \mathcal{O}(\varepsilon) \tag{5.133c}$$

$$\Delta \left(\xi^\beta \dot{z}^\gamma h^{R(1)}_{\beta\gamma} \right) = \xi^\beta \dot{z}^\gamma \nabla_\beta \zeta_\gamma + \dot{z}^\gamma \xi^\beta \nabla_\gamma \zeta_\beta + \mathcal{O}(\varepsilon) \tag{5.133d}$$

$$\Delta \left(\left(\frac{ds}{d\tau} \right) \mathcal{E} \dot{z}^\beta \dot{z}^\gamma h^{R(1)}_{\beta\gamma} \right) = 2 \left(\frac{ds}{d\tau} \right) \dot{z}^\beta \dot{z}^\gamma \nabla_\beta \zeta_\gamma + \mathcal{O}(\varepsilon) \tag{5.133e}$$

Now, we're prepared to perform the replacements. First, keeping all of the terms in order and

replacing back to the original velocities where convenient, the full expression is

$$\begin{aligned}\left\langle \Delta\frac{d\mathcal{E}}{d\tau}\right\rangle =&\frac{1}{2}\varepsilon\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}_\zeta g^{(0)}_{\mu\nu}(z(\tau))\right\rangle - \frac{1}{2}\varepsilon^2\left\langle u^\mu u^\nu(\mathcal{L}_\zeta+\pounds_\zeta)\mathcal{L}_\xi h^{R(1)}_{\mu\nu}\right\rangle\\ &-\frac{1}{2}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\zeta\mathcal{L}_\xi\mathcal{L}_\zeta g^{(0)}_{\mu\nu}\right\rangle - \frac{1}{4}\varepsilon^2\left\langle\left(\frac{ds}{d\tau}\right)^3 \dot z^\alpha \dot z^\beta \mathcal{L}_\zeta g^{(0)}_{\alpha\beta}\dot z^\mu \dot z^\nu\left(\mathcal{L}_\xi h_{\mu\nu}+\mathcal{L}_\xi\mathcal{L}_\zeta g^{(0)}_{\mu\nu}\right)\right\rangle_s\\ &+\frac{1}{4}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}_\zeta g^{(0)}_{\mu\nu}u^\alpha u^\beta \mathcal{L}_\zeta g^{(0)}_{\alpha\beta}\right\rangle + \frac{1}{4}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}_\zeta g_{\mu\nu}u^\alpha u^\beta h^{R(1)}_{\alpha\beta}\right\rangle\\ &+\frac{1}{4}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi h^{(1)}_{\mu\nu}u^\alpha u^\beta \mathcal{L}_\zeta g^{(0)}_{\alpha\beta}\right\rangle + \frac{1}{4}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}^2_\zeta g^{(0)}_{\mu\nu}\right\rangle\\ &+\frac{1}{2}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi(\mathcal{L}_\zeta+\pounds_\zeta)h^{R(1)}_{\mu\nu}\right\rangle - \varepsilon^2\partial_{\tilde\tau}\left\langle \xi^\beta u^\gamma(\nabla_\beta\zeta_\gamma+\nabla_\gamma\zeta_\beta)\right\rangle\\ &-\varepsilon^2\partial_{\tilde\tau}\left\langle u^\alpha u^\beta\nabla_\alpha\zeta_\beta\right\rangle+\mathcal{O}(\varepsilon^3).\end{aligned} \tag{5.134}$$

We now see that the first term on the second line cancels with the second term on the second line and the second term on the third line. The final term can be immediately related to a total time derivative plus an $\mathcal{O}(\varepsilon^3)$ acceleration term. We are left with (changing for brevity back to averages over tau):

$$\begin{aligned}\left\langle \Delta\frac{d\mathcal{E}}{d\tau}\right\rangle =&\frac{1}{2}\varepsilon\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}_\zeta g^{(0)}_{\mu\nu}(z(\tau))\right\rangle - \frac{1}{2}\varepsilon^2\left\langle (u^\mu u^\nu(\mathcal{L}_\zeta+\pounds_\zeta)\mathcal{L}_\xi h^{R(1)}_{\mu\nu}(z(\tau);z))\right\rangle\\ &-\frac{1}{2}\varepsilon^2\left\langle (u^\mu u^\nu \mathcal{L}_\zeta\mathcal{L}_\xi\mathcal{L}_\zeta g^{(0)}_{\mu\nu}(z(\tau)))\right\rangle + \frac{1}{4}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}_\zeta g_{\mu\nu}(z(\tau))u^\alpha u^\beta h^{R(1)}_{\alpha\beta}(z(\tau);z)\right\rangle\\ &+\frac{1}{4}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}^2_\zeta g^{(0)}_{\mu\nu}(z(\tau))\right\rangle + \frac{1}{2}\varepsilon^2\left\langle u^\mu u^\nu \mathcal{L}_\xi(\mathcal{L}_\zeta+\pounds_\zeta)h^{R(1)}_{\mu\nu}(z(\tau);z)\right\rangle\\ &-\varepsilon^2\partial_{\tilde\tau}\left\langle \xi^\beta u^\gamma(\nabla_\beta\zeta_\gamma+\nabla_\gamma\zeta_\beta)\right\rangle+\mathcal{O}(\varepsilon^3).\end{aligned} \tag{5.135}$$

We can immediately cancel the $\pounds$ derivatives Consider now the first term of this expansion. We may re-write this term as

$$\begin{aligned}\frac{1}{2}\varepsilon\left\langle u^\mu u^\nu \mathcal{L}_\xi\mathcal{L}_\zeta g^{(0)}_{\mu\nu}\right\rangle =&\varepsilon\left\langle u^\mu u^\nu \mathcal{L}_\xi\nabla_\mu\zeta_\nu\right\rangle\\ =&-\varepsilon^2\left\langle f^\mu\mathcal{L}_\xi\zeta_\mu\right\rangle+\varepsilon^2\partial_{\tilde\tau}\left\langle u^\nu\xi^\mu\nabla_\mu\zeta_\nu\right\rangle+\varepsilon^2\partial_{\tilde\tau}\left\langle u^\nu\zeta_\mu\nabla_\nu\xi^\mu\right\rangle,\end{aligned} \tag{5.136}$$

where we write $a^\mu=\varepsilon f^\mu+\mathcal{O}(\varepsilon^2)$. Note that the last term of (5.135) added to the second and third terms of (5.136) gives

$$-\varepsilon^2\partial_{\tilde\tau}\left\langle \xi^\beta u^\gamma(\nabla_\beta\zeta_\gamma+\nabla_\gamma\zeta_\beta)\right\rangle+\varepsilon^2\partial_{\tilde\tau}\left\langle u^\nu\xi^\mu\nabla_\mu\zeta_\nu\right\rangle+\varepsilon^2\partial_{\tilde\tau}\left\langle u^\nu\zeta_\mu\nabla_\nu\xi^\mu\right\rangle=2\varepsilon^2\partial_{\tilde\tau}\left\langle u^\nu\zeta_\mu\nabla_\nu\xi^\mu\right\rangle, \tag{5.137}$$

so we will need to manipulate the argument of the averaging operation a fair amount. Consider

$$\begin{aligned}\left\langle \zeta^\beta \dot{z}^\alpha \nabla_\alpha \xi_\beta \right\rangle_s &= -\left\langle \xi^\beta \dot{z}^\alpha \nabla_\alpha \zeta_\beta \right\rangle_s + \mathcal{O}(\varepsilon) \\ &= \left\langle \zeta^\gamma \dot{z}^\beta \nabla_\gamma \xi_\beta \right\rangle - \left\langle \zeta^\gamma \nabla_\gamma \left(\dot{z}^\beta \xi_\beta \right) \right\rangle + \mathcal{O}(\varepsilon) \\ &= -\left\langle \zeta^\beta \dot{z}^\gamma \nabla_\gamma \xi_\beta \right\rangle + \langle \mathcal{E} u^\mu u^\nu \nabla_\mu \zeta_\nu \rangle + \mathcal{O}(\varepsilon) \\ &= -\left\langle \zeta^\beta \dot{z}^\gamma \nabla_\gamma \xi_\beta \right\rangle + \mathcal{O}(\varepsilon) \\ \Rightarrow \left\langle \zeta^\beta \dot{z}^\gamma \nabla_\gamma \xi_\beta \right\rangle &= \mathcal{O}(\varepsilon) \end{aligned} \tag{5.138}$$

Subsituting all of these results back into (5.135) and rearranging the resulting expression, we obtain,

$$\begin{aligned}\left\langle \Delta \frac{d\mathcal{E}}{d\tau} \right\rangle = &-\varepsilon^2 \langle f^\mu \mathcal{L}_\xi \zeta_\mu \rangle - \frac{1}{2}\varepsilon^2 \left\langle u^\mu u^\nu [\mathcal{L}_\zeta, \mathcal{L}_\xi] h^{R(1)}_{\mu\nu} \right\rangle + \frac{1}{4}\varepsilon^2 \left\langle u^\mu u^\nu \mathcal{L}_\xi \mathcal{L}_\zeta g_{\mu\nu} u^\alpha u^\beta h^{R(1)}_{\alpha\beta} \right\rangle \\ &- \frac{1}{2}\varepsilon^2 \left\langle u^\mu u^\nu \left(\mathcal{L}_\zeta \mathcal{L}_\xi \mathcal{L}_\zeta - \frac{1}{2}\mathcal{L}_\xi \mathcal{L}_\zeta^2 \right) g^{(0)}_{\mu\nu} \right\rangle + \mathcal{O}(\varepsilon^3) \end{aligned} \tag{5.139}$$

We'll need to simplify this expression in steps. First, expand out the final term:

$$\begin{aligned}\frac{1}{2}\left\langle u^\mu u^\nu \mathcal{L}_\zeta \mathcal{L}_\xi \mathcal{L}_\zeta g^{(0)}_{\mu\nu} \right\rangle &= \langle u^\mu u^\nu \zeta^\gamma \nabla_\gamma \nabla_\mu \mathcal{L}_\xi \zeta_\nu \rangle + \langle u^\mu u^\nu \nabla_\mu \mathcal{L}_\xi \zeta^\gamma \nabla_\nu \zeta_\gamma \rangle + \langle u^\mu u^\nu \nabla_\gamma \mathcal{L}_\xi \zeta_\mu \nabla_\nu \zeta^\gamma \rangle \\ &= -\left\langle u^\alpha u^\beta \zeta^\gamma R_{\beta\delta\alpha\gamma} \mathcal{L}_\xi \zeta^\delta \right\rangle - \left\langle u^\alpha u^\beta \mathcal{L}_\xi \zeta^\gamma \nabla_\alpha \nabla_\beta \zeta_\gamma \right\rangle \end{aligned} \tag{5.140a}$$

$$\begin{aligned}\frac{1}{4}\left\langle u^\mu u^\nu \mathcal{L}_\xi \mathcal{L}_\zeta^2 g^{(0)}_{\mu\nu} \right\rangle &= \frac{1}{2}\left\langle u^\alpha u^\beta \mathcal{L}_\xi \left(\zeta^\gamma \nabla_\gamma \nabla_\alpha \zeta_\beta + \nabla_\alpha \zeta^\gamma \nabla_\gamma \zeta_\beta + \nabla_\alpha \zeta_\gamma \nabla_\beta \zeta^\gamma \right) \right\rangle \\ &= -\left\langle u^\alpha u^\beta \zeta^\gamma R_{\beta\delta\alpha\gamma} \mathcal{L}_\xi \zeta^\delta \right\rangle - \left\langle u^\alpha u^\beta \mathcal{L}_\xi \zeta^\gamma \nabla_\alpha \nabla_\beta \zeta_\gamma \right\rangle, \end{aligned} \tag{5.140b}$$

so the final term in (5.139) vanishes. The remaining terms expand as

$$-\varepsilon^2 \langle f^\mu \mathcal{L}_\xi \zeta_\mu \rangle = \frac{\varepsilon^2}{2}\left\langle \mathcal{L}_\xi \zeta^\beta u_\beta u^\gamma u^\delta u^\alpha \nabla_\delta h^{R(1)}_{\gamma\alpha} + 2\mathcal{L}_\xi \zeta^\beta u^\alpha u^\gamma \nabla_\alpha h^{R(1)}_{\beta\gamma} - \mathcal{L}_\xi \zeta^\beta u^\alpha u^\gamma \nabla_\beta h^{R(1)}_{\alpha\gamma} \right\rangle \tag{5.141a}$$

$$-\frac{1}{2}\varepsilon^2 \left\langle u^\mu u^\nu [\mathcal{L}_\zeta, \mathcal{L}_\xi] h^{R(1)}_{\mu\nu} \right\rangle = -\left\langle u^\alpha u^\beta \mathcal{L}_\xi \zeta^\gamma \nabla_\alpha h^{R(1)}_{\gamma\beta} \right\rangle + \frac{1}{2}\left\langle u^\alpha u^\beta \mathcal{L}_\xi \zeta^\gamma \nabla_\gamma h_{\alpha\beta} \right\rangle \tag{5.141b}$$

$$\frac{1}{4}\varepsilon^2 \left\langle u^\mu u^\nu \mathcal{L}_\xi \mathcal{L}_\zeta g^{(0)}_{\mu\nu} u^\alpha u^\beta h^{R(1)}_{\alpha\beta} \right\rangle = -\frac{1}{2}\left\langle u^\mu u^\nu \mathcal{L}_\xi \zeta_\mu u^\alpha u^\beta \nabla_\nu h^{R(1)}_{\alpha\beta} \right\rangle. \tag{5.141c}$$

We see now that these remaining three terms sum to zero. Therefore, the energy and angular momentum fluxes are gauge invariant through second order

$$\left\langle \Delta \frac{d\mathcal{E}}{d\tau} \right\rangle = \mathcal{O}(\varepsilon^3). \tag{5.142}$$

5.C Definitions and conventions for the Newman-Penrose formalism

There are a number of sign conventions when assembling the Newman-Penrose formalism. To avoid possible ambiguity, we list in this appendix all quantities relevant to the second order Teukolsky-Lousto-Campanelli wave equations discussed in Section 5.3. We follow all sign conventions of [360] (consistent with a signature $(-,+,+,+)$), further details of the Newman-Penrose formalism can be found in that paper, and a more in-depth discussion can be found in [382].

We use a null tetrad, $\{l^\mu, n^\mu, m^\mu, \bar{m}^\mu\} = \{e_1^\mu, e_2^\mu, e_3^\mu, e_4^\mu\}$ such that $l^\mu n_\mu = -1$, $m^\mu \bar{m}_\mu = 1$, and all other pairs of null tetrads are orthogonal. The directional derivatives along each of the tetrads are denoted $D = l^\mu \nabla_\mu$,$\Delta = n^\mu \nabla_\mu$,$\delta = m^\mu \nabla_\mu$, and $\bar{\delta} = \bar{m}^\mu \nabla_\mu$. The orthogonality conditions of the tetrads imply that we can write the background metric as,

$$g_{\mu\nu} = -l_\mu n_\nu + m_\mu \bar{m}_\nu \tag{5.143}$$

This metric form may be perturbed and inverted to obtain a construction for the first order tetrads in terms of the first order metric perturbation. Define the tetrad expansion as

$$e_a^\mu = e_a^{(0)\mu} + \varepsilon e_a^{(1)\mu} + \mathcal{O}(\varepsilon^2) \tag{5.144}$$

The metric perturbation is,

$$\begin{aligned} h^{(1)}_{\mu\nu} =& -2l^{(1)}{}_{(\mu} n_{\nu)} - 2n^{(1)}{}_{(\mu} l_{\nu)} + 2m^{(1)}{}_{(\mu} \bar{m}_{\nu)} + 2\bar{m}^{(1)}{}_{(\mu} m_{\nu)} \\ &+ 2m_{(\nu} \bar{m}^\lambda h^{(1)}_{\mu)\lambda} + 2\bar{m}_{(\nu} m^\lambda h^{(1)}_{\mu)\lambda} - 2l_{(\nu} n^\lambda h^{(1)}_{\mu)\lambda} - 2n_{(\nu} l^\lambda h^{(1)}_{\mu)\lambda}. \end{aligned} \tag{5.145}$$

These resulting tetrads are given also in [356], reproduced from [357] (up to sign convention),

$$l^{(1)\mu} = \frac{1}{2} h_{ll} n^\mu \tag{5.146a}$$

$$n^{(1)\mu} = \frac{1}{2} h_{nn} l^\mu + h_{nl} n^\mu \tag{5.146b}$$

$$m^{(1)\mu} = -\frac{1}{2} h_{mm} \bar{m}^\mu - \frac{1}{2} h_{m\bar{m}} m^\mu + h_{ml} n^\mu + h_{mn} l^\mu \tag{5.146c}$$

$$\bar{m}^{(1)\mu} = -\frac{1}{2} h_{\bar{m}\bar{m}} m^\mu - \frac{1}{2} h_{m\bar{m}} \bar{m}^\mu + h_{\bar{m}l} n^\mu + h_{\bar{m}n} l^\mu \tag{5.146d}$$

The full information contained in the connection can be expressed in terms of the spin coefficients, defined as

$$\kappa = -\, m^\mu l_{\mu;\nu} l^\nu \qquad \pi = -n^\mu \bar{m}_{\mu;\nu} l^\nu \qquad \varepsilon = -\frac{1}{2}\left(n^\mu l_{\mu;\nu} l^\nu + m^\mu \bar{m}_{\mu;\nu} l^\nu\right) \tag{5.147a}$$

$$\tau = -\, m^\mu l_{\mu;\nu} n^\nu \qquad \nu = -n^\mu \bar{m}_{\mu;\nu} n^\nu \qquad \gamma = -\frac{1}{2}\left(n^\mu l_{\mu;\nu} n^\nu + m^\mu \bar{m}_{\mu;\nu} n^\mu\right) \tag{5.147b}$$

$$\sigma = -\, m^\mu l_{\mu;\nu} m^\nu \qquad \mu = -n^\mu \bar{m}_{\mu;\nu} m^\nu \qquad \beta = -\frac{1}{2}\left(n^\mu l_{\mu;\nu} m^\nu + m^\mu \bar{m}_{\mu;\nu} m^\nu\right) \tag{5.147c}$$

$$\rho = -\, m^\mu l_{\mu;\nu} \bar{m}^\nu \qquad \lambda = -n^\mu \bar{m}_{\mu;\nu} \bar{m}^\nu \qquad \alpha = -\frac{1}{2}\left(n^\mu l_{\mu;\nu} \bar{m}^\nu + m^\mu \bar{m}_{\mu;\nu} \bar{m}^\nu\right) \tag{5.147d}$$

We were unable to fully reproduce the first order spin coefficients quoted in [356]. We have performed a full rederivation and insisted consistency with the Bianchi identities used in the original derivation. We have found a slightly corrected full set of first order spin coefficients,

$$\begin{aligned}\kappa^{(1)} =& -\kappa h_{ln} - \tfrac{1}{2}\kappa h_{m\bar{m}} - \tfrac{1}{2}\bar{\kappa} h_{mm} - (D - 2\varepsilon - \bar{\rho}) h_{lm} \\ &+ \sigma h_{l\bar{m}} - (\bar{\alpha} + \beta - \tfrac{1}{2}\bar{\pi} - \tfrac{1}{2}\tau - \tfrac{1}{2}\delta) h_{ll}\end{aligned} \tag{5.148a}$$

$$\sigma^{(1)} = -\tfrac{1}{2}\bar{\lambda} h_{ll} - (\tfrac{1}{2}D - \varepsilon + \bar{\varepsilon} + \tfrac{1}{2}\rho - \tfrac{1}{2}\bar{\rho}) h_{mm} - (-\bar{\pi} - \tau) h_{lm} \tag{5.148b}$$

$$\begin{aligned}\nu^{(1)} =& \lambda h_{nm} - (-\Delta - 2\gamma - \bar{\mu}) h_{n\bar{m}} + \nu h_{ln} - \tfrac{1}{2}\nu h_{m\bar{m}} \\ &- \tfrac{1}{2}\bar{\nu} h_{\bar{m}\bar{m}} - (\alpha + \bar{\beta} - \tfrac{1}{2}\pi - \tfrac{1}{2}\bar{\tau} + \tfrac{1}{2}\bar{\delta}) h_{nn}\end{aligned} \tag{5.148c}$$

$$\lambda^{(1)} = \lambda h_{ln} - (-\tfrac{1}{2}\Delta - \gamma + \bar{\gamma} + \tfrac{1}{2}\mu - \tfrac{1}{2}\bar{\mu}) h_{\bar{m}\bar{m}} - \tfrac{1}{2}\bar{\sigma} h_{nn} - (-\pi - \bar{\tau}) h_{n\bar{m}} \tag{5.148d}$$

$$\begin{aligned}\mu^{(1)} =& -(-\tfrac{1}{2}\mu - \tfrac{1}{2}\bar{\mu}) h_{ln} - (-\tfrac{1}{2}\Delta + \tfrac{1}{2}\mu - \tfrac{1}{2}\bar{\mu}) h_{m\bar{m}} - (-\tfrac{1}{2}\delta - \beta - \tfrac{1}{2}\tau) h_{n\bar{m}} \\ &- (\tfrac{1}{2}\bar{\delta} + \bar{\beta} - \pi - \tfrac{1}{2}\bar{\tau}) h_{nm} + \tfrac{1}{2}\nu h_{lm} - \tfrac{1}{2}\bar{\nu} h_{l\bar{m}} - \tfrac{1}{2}\rho h_{nn}\end{aligned} \tag{5.148e}$$

$$\begin{aligned}\rho^{(1)} =& \tfrac{1}{2}\kappa h_{n\bar{m}} - \tfrac{1}{2}\bar{\kappa} h_{nm} - \tfrac{1}{2}\mu h_{ll} - (\tfrac{1}{2}\bar{\delta} - \alpha - \tfrac{1}{2}\pi) h_{lm} - (\tfrac{1}{2}\rho - \tfrac{1}{2}\bar{\rho}) h_{ln} \\ &- (\tfrac{1}{2}D + \tfrac{1}{2}\rho + \tfrac{1}{2}\bar{\rho}) h_{m\bar{m}} - (-\tfrac{1}{2}\delta + \bar{\alpha} - \tfrac{1}{2}\bar{\pi} - \tau) h_{l\bar{m}}\end{aligned} \tag{5.148f}$$

$$\begin{aligned}\varepsilon^{(1)} =& \tfrac{1}{4}\kappa h_{n\bar{m}} - \tfrac{1}{4}\bar{\kappa} h_{nm} - (-\tfrac{1}{4}\Delta + \tfrac{1}{2}\bar{\gamma} + \tfrac{1}{4}\mu - \tfrac{1}{4}\bar{\mu}) h_{ll} - (\tfrac{1}{2}D + \tfrac{1}{4}\rho - \tfrac{1}{4}\bar{\rho}) h_{ln} \\ &- (\tfrac{1}{4}\rho - \tfrac{1}{4}\bar{\rho}) h_{m\bar{m}} + \tfrac{1}{4}\sigma h_{\bar{m}\bar{m}} - \tfrac{1}{4}\bar{\sigma} h_{mm} - (-\tfrac{1}{4}\delta + \tfrac{1}{2}\bar{\alpha} - \tfrac{1}{4}\bar{\pi} - \tfrac{1}{2}\tau) h_{l\bar{m}} \\ &- (\tfrac{1}{4}\bar{\delta} - \tfrac{1}{2}\alpha - \tfrac{3}{4}\pi - \tfrac{1}{2}\bar{\tau}) h_{lm}\end{aligned} \tag{5.148g}$$

$$\begin{aligned}\pi^{(1)} =& \lambda \tfrac{1}{2} h_{lm} - (-\tfrac{1}{2}\Delta + \bar{\gamma} - \tfrac{1}{2}\bar{\mu}) h_{l\bar{m}} - (-\tfrac{1}{2}D - \varepsilon + \tfrac{1}{2}\rho) h_{n\bar{m}} - \tfrac{1}{2}\bar{\sigma} h_{nm} \\ &+ \tfrac{1}{2}\tau h_{\bar{m}\bar{m}} - (\tfrac{1}{2}\bar{\delta} + \tfrac{1}{2}\pi + \tfrac{1}{2}\bar{\tau}) h_{ln} + \tfrac{1}{2}\bar{\tau} h_{m\bar{m}}\end{aligned} \tag{5.148h}$$

$$\begin{aligned}\tau^{(1)} =& -\tfrac{1}{2}\bar{\lambda}h_{l\bar{m}} - (\tfrac{1}{2}\Delta - \gamma + \tfrac{1}{2}\mu)h_{lm} + \tfrac{1}{2}\pi h_{mm} + \tfrac{1}{2}\bar{\pi}h_{m\bar{m}} \\ & - (\tfrac{1}{2}D + \bar{\varepsilon} - \tfrac{1}{2}\bar{\rho})h_{nm} + \tfrac{1}{2}\sigma h_{n\bar{m}} - (-\tfrac{1}{2}\delta - \tfrac{1}{2}\bar{\pi} - \tfrac{1}{2}\tau)h_{ln} \qquad (5.148\text{i}) \\ \alpha^{(1)} =& -\tfrac{1}{4}\bar{\kappa}h_{nn} + \tfrac{3}{4}\lambda h_{lm} - (-\tfrac{1}{4}\Delta - \gamma + \tfrac{1}{2}\bar{\gamma} + \tfrac{1}{2}\mu - \tfrac{1}{4}\bar{\mu})h_{l\bar{m}} - \tfrac{1}{4}\nu h_{ll} \\ & - (\tfrac{1}{4}D - \tfrac{1}{2}\varepsilon + \tfrac{1}{4}\rho + \tfrac{1}{2}\bar{\rho})h_{n\bar{m}} - \tfrac{1}{4}\bar{\sigma}h_{nm} - (-\tfrac{1}{4}\delta + \tfrac{1}{2}\bar{\alpha} - \tfrac{1}{4}\bar{\pi} - \tfrac{1}{4}\tau)h_{\bar{m}\bar{m}} \\ & - (\tfrac{1}{4}\bar{\delta} - \tfrac{1}{4}\pi - \tfrac{1}{4}\bar{\tau})h_{ln} - (\tfrac{1}{4}\bar{\delta} + \tfrac{1}{2}\alpha - \tfrac{1}{4}\pi - \tfrac{1}{4}\bar{\tau})h_{m\bar{m}} \qquad (5.148\text{j}) \\ \beta^{(1)} =& -\tfrac{1}{4}\kappa h_{nn} - \tfrac{1}{4}\bar{\lambda}h_{l\bar{m}} - (-\tfrac{1}{4}\Delta - \tfrac{1}{2}\gamma - \tfrac{1}{4}\mu - \tfrac{1}{2}\bar{\mu})h_{lm} - \tfrac{1}{4}\bar{\nu}h_{ll} \\ & - (\tfrac{1}{4}D - \varepsilon + \tfrac{1}{2}\bar{\varepsilon} - \tfrac{1}{2}\rho + \tfrac{1}{4}\bar{\rho})h_{nm} + \tfrac{3}{4}\sigma h_{n\bar{m}} - (\tfrac{1}{4}\delta - \tfrac{1}{4}\bar{\pi} - \tfrac{1}{4}\tau)h_{ln} \\ & - (-\tfrac{1}{4}\delta + \tfrac{1}{2}\beta - \tfrac{1}{4}\bar{\pi} - \tfrac{1}{4}\tau)h_{m\bar{m}} - (\tfrac{1}{4}\bar{\delta} + \tfrac{1}{2}\bar{\beta} - \tfrac{1}{4}\pi - \tfrac{1}{4}\bar{\tau})h_{mm} \qquad (5.148\text{k}) \\ \gamma^{(1)} =& -\tfrac{1}{4}\lambda h_{mm} + \tfrac{1}{4}\bar{\lambda}h_{\bar{m}\bar{m}} + (\tfrac{1}{4}\mu - \tfrac{1}{4}\bar{\mu})h_{m\bar{m}} + (-\gamma + \tfrac{1}{4}\mu - \tfrac{1}{4}\bar{\mu})h_{ln} \\ & - \tfrac{1}{4}\nu h_{lm} + \tfrac{1}{4}\bar{\nu}h_{l\bar{m}} + (\tfrac{1}{4}D + \tfrac{1}{2}\bar{\varepsilon} + \tfrac{1}{4}\rho - \tfrac{1}{4}\bar{\rho})h_{nn} + (-\tfrac{1}{4}\delta - \tfrac{1}{2}\beta - \tfrac{1}{2}\bar{\pi} - \tfrac{3}{4}\tau)h_{n\bar{m}} \\ & + (\tfrac{1}{4}\bar{\delta} + \tfrac{1}{2}\bar{\beta} - \tfrac{1}{2}\pi - \tfrac{1}{4}\bar{\tau})h_{nm} \qquad (5.148\text{l})\end{aligned}$$

The Weyl scalars which appear in the Teukolsky-Lousto-Campanelli wave equations are defined as tetrad components of the Weyl tensor,

$$\begin{aligned}\psi_0 =& C_{\alpha\beta\gamma\delta}l^\alpha m^\beta l^\gamma m^\delta, \qquad (5.149\text{a}) \\ \psi_1 =& C_{\alpha\beta\gamma\delta}l^\alpha m^\beta l^\gamma n^\delta, \qquad (5.149\text{b}) \\ \psi_2 =& C_{\alpha\beta\gamma\delta}l^\alpha m^\beta \bar{m}^\gamma n^\delta, \qquad (5.149\text{c}) \\ \psi_3 =& C_{\alpha\beta\gamma\delta}l^\alpha n^\beta \bar{m}^\gamma n^\delta, \qquad (5.149\text{d}) \\ \psi_4 =& C_{\alpha\beta\gamma\delta}n^\alpha \bar{m}^\beta n^\gamma \bar{m}^\delta, \qquad (5.149\text{e})\end{aligned}$$

which can then be expanded in terms of quadratic functions of the spin coefficients and their derivatives. The subleading Weyl scalars can be extracted by perturbing the identities extracted

from the Newman-Penrose equations [356, 382],

$$\psi_0^{(n)} = \sum_{p=0}^{n} (D - 3\varepsilon + \bar{\varepsilon} - \rho - \bar{\rho})^{(n-p)} \sigma^{(p)} - (\delta - \bar{\alpha} - 3\beta + \bar{\pi} - \tau)^{(n-p)} \kappa^{(p)} \tag{5.150a}$$

$$\psi_1^{(n)} = \sum_{p=0}^{n} (D + \bar{\varepsilon} - \bar{\rho})^{(n-p)} \beta^{(p)} - (\delta - \bar{\alpha} + \bar{\pi})^{(n-p)} \varepsilon^{(p)} - (\alpha + \pi)^{(n-p)} \sigma^{(p)} + (\gamma + \mu)^{(n-p)} \kappa^{(p)} \tag{5.150b}$$

$$\begin{aligned} \psi_2^{(n)} = \frac{1}{3} \sum_{p=0}^{n} \Big[& (\bar{\delta} - 2\alpha + \bar{\beta} - \pi - \bar{\tau})^{(n-p)} \beta^{(p)} - (\delta - \bar{\alpha} + \bar{\pi} + \tau)^{(n-p)} \alpha^{(p)} \\ & + (D + \varepsilon + \bar{\varepsilon} + \rho - \bar{\rho})^{(n-p)} \gamma^{(p)} - (\Delta - \bar{\gamma} - \gamma + \bar{\mu} - \mu)^{(n-p)} \varepsilon^{(p)} \\ & + (\bar{\delta} - \alpha + \bar{\beta} - \bar{\tau} - \pi)^{(n-p)} \tau^{(p)} + 2\nu^{(n-p)} \kappa^{(p)} - 2\lambda^{(n-p)} \sigma^{(p)} \\ & - (\Delta - \bar{\gamma} - \gamma + \bar{\mu} - \mu)^{(n-p)} \rho^{(p)} \Big] \end{aligned} \tag{5.150c}$$

$$\psi_3^{(n)} = \sum_{p=0}^{n} (\bar{\delta} + \bar{\beta} - \bar{\tau})^{(n-p)} \gamma^{(p)} - (\Delta - \bar{\gamma} + \bar{\mu})^{(n-p)} \alpha^{(p)} + (\varepsilon + \rho)^{(n-p)} \nu^{(p)} - (\beta + \tau)^{(n-p)} \lambda^{(p)} \tag{5.150d}$$

$$\psi_4^{(n)} = \sum_{p=0}^{n} (\bar{\delta} + 3\alpha + \bar{\beta} + \pi - \bar{\tau})^{(n-p)} \nu^{(p)} - (\Delta - \bar{\gamma} + 3\gamma + \mu + \bar{\mu})^{(n-p)} \lambda^{(p)} \tag{5.150e}$$

Chapter 5 Bibliography

[304] G. O. A. Team, *The ESA-L3 Gravitational Wave Mission* (2016)

[305] I. Mandel, D. A. Brown, J. R. Gair and M. C. Miller, *Rates and Characteristics of Intermediate-Mass-Ratio Inspirals Detectable by Advanced LIGO*, Astrophys. J. **681** (2008), pp. 1431–1447

[306] T. Hinderer and E. E. Flanagan, *Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime: Orbital motion*, Phys. Rev. D **78** (2008), p. 064028,
URL: http://link.aps.org/doi/10.1103/PhysRevD.78.064028

[307] E. E. Flanagan and T. Hinderer, *Transient resonances in the inspirals of point particles into black holes*, Phys. Rev. Lett. **109** (2012), p. 071102

[308] L. Barack and N. Sago, *Gravitational self-force on a particle in eccentric orbit around a Schwarzschild black hole*, Phys. Rev. D **81** (2010), p. 084021

[309] S. Akcay, N. Warburton and L. Barack, *Frequency-domain algorithm for the Lorenz-gauge gravitational self-force*, Phys. Rev. D **88** (2013) (10), p. 104009

[310] B. Wardell, C. R. Galley, A. Zenginoglu, M. Casals, S. R. Dolan *et al.*, *Self-force via Green functions and worldline integration*, Phys. Rev. D **89** (2014), p. 084021

[311] T. Osburn, E. Forseth, C. R. Evans and S. Hopper, *Lorenz gauge gravitational self-force calculations of eccentric binaries using a frequency domain procedure*, Phys. Rev. **D90** (2014) (10), p. 104031

[312] B. Wardell, *Self-force: Computational Strategies*, Fund. Theor. Phys. **179** (2015), pp. 487–522

[313] M. van de Meent, *Gravitational self-force on eccentric equatorial orbits around a Kerr black hole*, Phys. Rev. **D94** (2016) (4), p. 044034

[314] N. K. Johnson-McDaniel, A. G. Shah and B. F. Whiting, *Experimental mathematics meets gravitational self-force*, Phys. Rev. **D92** (2015) (4), p. 044007

[315] S. L. Detweiler, *A consequence of the gravitational self-force for circular orbits of the Schwarzschild geometry*, Phys. Rev. D **77** (2008), p. 124026

[316] L. Barack and N. Sago, *Beyond the geodesic approximation: conservative effects of the gravitational self-force in eccentric orbits around a Schwarzschild black hole*, Phys. Rev. D **83** (2011), p. 084023

[317] S. R. Dolan, P. Nolan, A. C. Ottewill, N. Warburton and B. Wardell, *Tidal invariants for compact binaries on quasi-circular orbits* (2014)

[318] S. Akcay, D. Dempsey and S. Dolan, *Spin-orbit precession for eccentric black hole binaries at first order in the mass ratio* (2016)

[319] M. van de Meent, *Self-force corrections to the periapsis advance around a spinning black hole* (2016)

[320] L. Barack and A. Ori, *Mode sum regularization approach for the selfforce in black hole space-time*, Phys. Rev. **D61** (2000), p. 061502

[321] L. Barack, Y. Mino, H. Nakano, A. Ori and M. Sasaki, *Calculating the gravitational selfforce in Schwarzschild space-time*, Phys. Rev. Lett. **88** (2002), p. 091101

[322] L. Barack and D. A. Golbourn, *Scalar-field perturbations from a particle orbiting a black hole using numerical evolution in 2+1 dimensions*, Phys. Rev. **D76** (2007), p. 044020

[323] I. Vega and S. L. Detweiler, *Regularization of fields for self-force problems in curved spacetime: Foundations and a time-domain application*, Phys. Rev. **D77** (2008), p. 084008

[324] Y. Mino, *Perturbative approach to an orbital evolution around a supermassive black hole*, Phys. Rev. D **67** (2003), p. 084027

[325] N. Sago, T. Tanaka, W. Hikida, K. Ganz and H. Nakano, *The Adiabatic evolution of orbital parameters in the Kerr spacetime*, Prog. Theor. Phys. **115** (2006), pp. 873–907

[326] S. Drasco, E. E. Flanagan and S. A. Hughes, *Computing inspirals in Kerr in the adiabatic regime. I. The Scalar case*, Class. Quant. Grav. **22** (2005), pp. S801–846

[327] R. Fujita, W. Hikida and H. Tagoshi, *An Efficient Numerical Method for Computing Gravitational Waves Induced by a Particle Moving on Eccentric Inclined Orbits around a Kerr Black Hole*, Prog. Theor. Phys. **121** (2009), pp. 843–874

[328] S. A. Hughes, *Adiabatic and post-adiabatic approaches to extreme mass ratio inspiral* (2016), URL: https://inspirehep.net/record/1414738/files/arXiv:1601.02042.pdf

[329] A. Pound and E. Poisson, *Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals*, Phys. Rev. **D77** (2008), p. 044013

[330] J. R. Gair, E. E. Flanagan, S. Drasco, T. Hinderer and S. Babak, *Forced motion near black holes*, Phys. Rev. D **83** (2011), p. 044037

[331] N. Warburton, S. Akcay, L. Barack, J. R. Gair and N. Sago, *Evolution of inspiral orbits around a Schwarzschild black hole*, Phys. Rev. D **85** (2012), p. 061501

[332] T. Osburn, N. Warburton and C. R. Evans, *Highly eccentric inspirals into a black hole*, Phys. Rev. **D93** (2016) (6), p. 064024

[333] A. Pound, *Self-consistent gravitational self-force*, Phys. Rev. **D81** (2010), p. 024023

[334] P. Diener, I. Vega, B. Wardell and S. Detweiler, *Self-consistent orbital evolution of a particle around a Schwarzschild black hole*, Phys. Rev. Lett. **108** (2012), p. 191102

[335] A. Spallicci, P. Ritter, S. Jubertie, S. Cordier and S. Aoudia, *Towards a self-consistent orbital evolution for EMRIs*, ASP Conf. Ser. **467** (2013), p. 221

[336] R. Ruffini and M. Sasaki, *On a Semirelativistic Treatment of the Gravitational Radiation From a Mass Thrusted Into a Black Hole*, Prog. Theor. Phys. **66** (1981), pp. 1627–1638

[337] S. Babak, H. Fang, J. R. Gair, K. Glampedakis and S. A. Hughes, *'Kludge' gravitational waveforms for a test-body orbiting a Kerr black hole*, Phys. Rev. **D75** (2007), p. 024005, [Erratum: Phys. Rev.D77,04990(2008)]

[338] S. A. Hughes, S. Drasco, E. E. Flanagan and J. Franklin, *Gravitational radiation reaction and inspiral waveforms in the adiabatic limit*, Phys. Rev. Lett. **94** (2005), p. 221101

[339] N. Warburton, S. Akcay, L. Barack, J. R. Gair and N. Sago, *Evolution of inspiral orbits around a Schwarzschild black hole*, Phys. Rev. **D85** (2012), p. 061501

[340] M. Van De Meent and N. Warburton, *Fast Self-forced Inspirals* (2018)

[341] A. Pound and E. Poisson, *Multi-scale analysis of the electromagnetic self-force in a weak gravitational field*, Phys. Rev. **D77** (2008), p. 044012

[342] Y. Mino and R. Price, *Two-timescale adiabatic expansion of a scalar field model*, Phys. Rev. D **77** (2008), p. 064001

[343] A. Pound, *Second-order perturbation theory: problems on large scales*, Phys. Rev. **D92** (2015) (10), p. 104047

[344] E. E. Flanagan, T. Hinderer, J. Moxon and A. Pound, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. I. Foundations* ((In prep.))

[345] E. E. Flanagan, T. Hinderer, J. Moxon and A. Pound, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. III. Dynamics of the far zone waves* ((In prep.))

[346] S. Isoyama, A. Pound, T. Tanaka and K. Yamada, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. IV. Dynamics of the near-horizon region* ((In prep.))

[347] J. Brink, M. Geyer and T. Hinderer, *Orbital resonances around Black holes*, Phys. Rev. Lett. **114** (2015) (8), p. 081102

[348] S. Isoyama, R. Fujita, H. Nakano, N. Sago and T. Tanaka, *Evolution of the Carter constant for resonant inspirals into a Kerr black hole: I. The scalar case*, PTEP **2013** (2013) (6), p. 063E01

[349] C. P. L. Berry, R. H. Cole, P. CaÃśizares and J. R. Gair, *Importance of transient resonances in extreme-mass-ratio inspirals*, Phys. Rev. **D94** (2016) (12), p. 124042

[350] J. Kevorkian and J. D. Cole, *Multiple Scale and Singular Perturbation Methods* (Springer, New York, 1996)

[351] A. Pound, *Second-order gravitational self-force*, Phys. Rev. Lett. **109** (2012), p. 051101

[352] A. Pound, *Nonlinear gravitational self-force: second-order equation of motion* (2017)

[353] A. Pound and J. Miller, *Practical, covariant puncture for second-order self-force calculations*, Phys. Rev. **D89** (2014) (10), p. 104020

[354] S. E. Gralla, *Second Order Gravitational Self Force*, Phys. Rev. **D85** (2012), p. 124011

[355] S. A. Teukolsky, *Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations*, Astrophys. J. **185** (1973), pp. 635–647

[356] M. Campanelli and C. O. Lousto, *Second order gauge invariant gravitational perturbations of a Kerr black hole*, Phys. Rev. **D59** (1999), p. 124022

[357] P. L. Chrzanowski, *Vector Potential and Metric Perturbations of a Rotating Black Hole*, Phys. Rev. **D11** (1975), pp. 2042–2062

[358] J. M. Cohen and L. S. Kegeles, *Electromagnetic fields in curved spaces - a constructive procedure*, Phys. Rev. **D10** (1974), pp. 1070–1084

[359] A. Pound, C. Merlin and L. Barack, *Gravitational self-force from radiation-gauge metric perturbations*, Phys. Rev. **D89** (2014) (2), p. 024009

[360] C. Merlin, A. Ori, L. Barack, A. Pound and M. van de Meent, *Completion of metric reconstruction for a particle orbiting a Kerr black hole* (2016)

[361] M. van de Meent and A. G. Shah, *Metric perturbations produced by eccentric equatorial orbits around a Kerr black hole*, Phys. Rev. **D92** (2015) (6), p. 064025

[362] R. M. Wald, *Construction of Solutions of Gravitational, Electromagnetic, Or Other Perturbation Equations from Solutions of Decoupled Equations*, Phys. Rev. Lett. **41** (1978), pp. 203–206

[363] C. O. Lousto and B. F. Whiting, *Reconstruction of black hole metric perturbations from Weyl curvature*, Phys. Rev. **D66** (2002), p. 024026

[364] T. S. Keidl, J. L. Friedman and A. G. Wiseman, *On finding fields and self-force in a gauge appropriate to separable wave equations*, Phys. Rev. **D75** (2007), p. 124009

[365] A. Ori, *Reconstruction of inhomogeneous metric perturbations and electromagnetic four potential in Kerr space-time*, Phys. Rev. **D67** (2003), p. 124010

[366] W. J. Thompson, *Fourier series and the Gibbs phenomenon*, Am. J. Phys. **60** (1992), p. 425

[367] L. Barack, A. Ori and N. Sago, *Frequency-domain calculation of the self force: The High-frequency problem and its resolution*, Phys. Rev. **D78** (2008), p. 084021

[368] M. van De Meent, *The mass and angular momentum of reconstructed metric perturbations*, Class. Quant. Grav. **34** (2017) (12), p. 124003

[369] S. Hopper and C. R. Evans, *Gravitational perturbations and metric reconstruction: Method of extended homogeneous solutions applied to eccentric orbits on a Schwarzschild black hole*, Phys. Rev. **D82** (2010), p. 084010

[370] S. Hopper and C. R. Evans, *Metric perturbations from eccentric orbits on a Schwarzschild black hole: I. Odd-parity Regge-Wheeler to Lorenz gauge transformation and two new methods to circumvent the Gibbs phenomenon*, Phys. Rev. **D87** (2013) (6), p. 064008

[371] D. V. Gal'tsov, *Radiation reaction in the Kerr gravitational field*, Journal of Physics A: Mathematical and General **15** (1982) (12), p. 3737,
URL: http://stacks.iop.org/0305-4470/15/i=12/a=025

[372] N. Sago and R. Fujita, *Calculation of radiation reaction effect on orbital parameters in Kerr spacetime*, PTEP **2015** (2015) (7), p. 073E03

[373] K. Ganz, W. Hikida, H. Nakano, N. Sago and T. Tanaka, *Adiabatic Evolution of Three 'Constants' of Motion for Greatly Inclined Orbits in Kerr Spacetime*, Prog. Theor. Phys. **117** (2007), pp. 1041–1066

[374] D. Bini and T. Damour, *Gravitational radiation reaction along general orbits in the effective one-body formalism*, Phys. Rev. D **86** (2012), p. 124012,
URL: https://link.aps.org/doi/10.1103/PhysRevD.86.124012

[375] G. A. Schott, *Electromagnetic Radiation*, Philos. Mag. **29** (1915), p. 49

[376] J. R. Gair, E. E. Flanagan, S. Drasco, T. Hinderer and S. Babak, *Forced motion near black holes*, Phys. Rev. **D83** (2011), p. 044037

[377] Y. M. M. S. T. Tanaka, *Gravitational radiation reaction to a point particle*, Phys. Rev. D (1996), p. 55.3457

[378] T. C. Quinn and R. M. Wald, *Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime*, Phys. Rev. D **56** (1997), pp. 3381–3394,
URL: http://link.aps.org/doi/10.1103/PhysRevD.56.3381

[379] J. Vines and Ã. Ã. Flanagan, *Is motion under the conservative self-force in black hole spacetimes an integrable Hamiltonian system?*, Phys. Rev. **D92** (2015), p. 064039

[380] W. Schmidt, *Celestial mechanics in Kerr space-time*, Class. Quant. Grav. **19** (2002), p. 2743

[381] A. Pound, *Gauge and motion in perturbation theory*, Phys. Rev. **D92** (2015) (4), p. 044021

[382] J. M. Stewart, *Advanced general relativity* (Cambridge University Press, 1994)

Chapter 6

An overview of computations of the dynamics of Far Zone waves in the multiscale approximation framework and synthesis of multiscale approximations

Coauthors:

Éanna Flanagan, Cornell University

Tanja Hinderer, University of Maryland

Adam Pound, University of Southampton

6.1 | Overview

In this chapter, I present an overview of the computational techniques which will be necessary to fully compute the dynamics of the region distant from the inspiral system in the multiscale tapestry approximation discussed in Chapters 4 and 5. These computations are an important part of the ongoing effort to produce a comprehensive framework for working with the multiscale approximation for the full duration of an extreme mass ratio inspiral. In this chapter, we describe two different routes for the Far Zone computation, one which is based on a modification of Post-Minkowski theory, and one which is based on a modification of geometric optics.

The first expansion, which is discussed in Section 6.2, is based closely on Adam Pound's prior work examining the distant, weak-field expansion of a nonlinear scalar field [383], and uses techniques first presented by Blanchet and Damour [384, 385] for working with extended sources in Post-

Minkowski expansions. I do not present full details for the Post-Minkowski treatment of the Far Zone, and instead give only the general outline of the method to be presented in the publication in preparation [386]. The second method we have explored is an adapted version of a geometric optics expansion, which is presented in Section 6.3. The geometric optics expansion takes a noticeably different flavor to the Interaction Zone multiscale expansion, as the spatial scales are similar to the long timescale, and only a fast retarded time variable describes rapid phase evolution. I conclude by summarizing in Section 6.4 a detailed picture of our current understanding of the algorithm for computations of post-adiabatic waveforms in the multiscale tapestry approximation. In addition, I give an overview of the final computations necessary to complete our tapestry framework for the Einstein field equations in the high mass ratio case.

6.2 | Post-Minkowski computation

The goal of the Far Zone computation is to find the inhomogeneous solution to the wave equation, to accuracy $\mathcal{O}(\varepsilon^2)$, at the interface to the Interaction Zone $r \ll M/\varepsilon$ and to the very Far Zone $r \gg M/\varepsilon$. The metric perturbations vary in magnitude in the Far Zone via their dependence on r^{-1}, so it is important for understanding the scale of various quadratically sourced contributions to determine the power of r^{-1} at which they enter. For this discussion, it is useful to take the homogeneous solutions [384, 385] as a point of reference:

$$h^{00}_{(1)}[x^\mu] = -4\sum_{l\geq 0}\frac{(-1)^l}{l!}\partial_L\left[r^{-1}M_L(t-r)\right] \tag{6.1a}$$

$$h^{0i}_{(1)}[x^\mu] = 4\sum_{l\geq 0}\frac{(-1)^l}{l!}\partial_{L-1}\left[r^{-1}M'_{i\,L-1}(t-r)\right] + 4\sum_{l\geq 1}\frac{(-1)^l l}{(l+1)!}\varepsilon_{iab}\partial_{a\,L-1}\left[r^{-1}S_{b\,L-1}(t-r)\right] \tag{6.1b}$$

$$h^{ij}_{(1)}[x^\mu] = -4\sum_{l\geq 2}\frac{(-1)^l}{l!}\partial_{L-2}\left[r^{-1}M''_{ij\,L-2}(t-r)\right] - 8\sum_{l\geq 2}\frac{(-1)^l l}{(l+1)!}\partial_{a\,L-2}\left[r^{-1}\varepsilon_{ab(i}S'_{j)b\,L-2}(t-r)\right] \tag{6.1c}$$

In these equations, we follow the notation of [384, 385] in the use of the capital L to denote spatial multi-indices $i_1, i_2, i_3, \ldots i_l$, and the M and S denote electric-type and magnetic-type multipole moments of the homogeneous field.

Far from the Interaction Zone $r \gg M$, the leading terms $\sim r^{-1}$ in these expansions will dominate. Throughout the Far Zone, terms arising from the square of the first order metric perturbations will

form a leading contribution to the quadratic source. Using the Post-Minkowski computation, we seek to ascertain the effect at subleading order due to quadratic combinations of the leading homogeneous modes (6.1).

We consider a double expansion of the metric perturbation in the Far Zone ($r \gg M$) in powers of the central mass M and the mass ratio ε. Up to a trivial redefinition of the expansion parameters, this is identical to a double expansion in the two masses of the system μ and M.

$$\begin{aligned} g_{\mu\nu} \equiv \eta_{\mu\nu} &+ M h^{(0,1)}_{\mu\nu} + M^2 h^{(0,2)}_{\mu\nu} + M^3 h^{(0,3)}_{\mu\nu} + \mathcal{O}(M^4) \\ &+ \mu h^{(1,0)}_{\mu\nu}(\varphi, \tilde{u}) + M\mu h^{(1,1)}_{\mu\nu}(\varphi, \tilde{u}) + \mu M^2 h^{(1,2)}_{\mu\nu}(\varphi, \tilde{u}) + \mathcal{O}(\mu M^3) \\ &+ \mu^2 h^{(2,0)}_{\mu\nu}(\varphi, \tilde{u}) + M\mu^2 h^{(2,1)}_{\mu\nu}(\varphi, \tilde{u}) + \mathcal{O}(\mu^2 M^2) \\ &+ \mathcal{O}(\mu^3), \end{aligned} \tag{6.2}$$

where all terms $h^{(0,n)}$ are immediately determined by the expansion of the background Kerr or Schwarzschild metric. The first term at $\mathcal{O}(\varepsilon)$ arises also at $\mathcal{O}(M)$, as the matched homogeneous radiation solutions scale as $\mu/r = \varepsilon M/r$. I have explicitly notated the expected fast φ_A and slow $\tilde{u}$ retarded time dependence of each order of the metric perturbation, as the nature of time dependence is important for the scale of the solution arising from the nonlinear source in the Post-Minkowski expansion.

We expand the wave equation in positive powers of M and μ. The resulting set of differential equations are all of the form

$$\Box h^{(n,p)} = S^{(n,p)}, \tag{6.3}$$

so the resulting metric perturbations $h^{(n,p)}$ are inhomogeneous tensor solutions to the vacuum wave operator $\Box$. In the overview presentation below, we only develop the approximations to the integrals described in [384, 385] in the context of a scalar field, although little alteration should be required to apply the same techniques to tensor harmonics. We will consider the case in which the source may be written as $S = r^{-k}\hat{n}_L f(u)$. For such sources, the methods of [384,385] can be used to derive

the particular solution via

$$\begin{aligned}\Box_{\rm ret}^{-1}[r^{-k}\hat{n}_L F(\varphi(\tilde{u})/\varepsilon,\varepsilon u)] &= \mathrm{FP}\lim_{B\to 0}\Box_{\rm ret}^{-1}[r^{B-k}\hat{n}_L F(\varphi,\tilde{u})]\\ &= \mathrm{FP}\lim_{B\to 0}\frac{1}{K_k^B}\int_{-\infty}^{u} ds F(\varphi(s),\varepsilon s)\hat{\partial}_L\frac{(u-s)^{B-k+l+2}-(v-s)^{B-k+l+2}}{r}\end{aligned} \tag{6.4}$$

where

$$K_k^B = 2^{B-k+3}\prod_{i=0}^{l}(B-k+2-i), \tag{6.5}$$

and FP denotes the operation of taking only the finite part in the limit of $B\to 0$, removing any poles. The complete details of this form of the Post-Minkowski solution and the justification of the Finite Part using analytic continuation is given in [384, 385].

We discuss this particular solution separately as the sum of a homogeneous solution $\hat{\partial}_L$ and a new particular solution,

$$\Box_{\rm ret}^{-1}[r^{B-k}\hat{n}_L F(\varphi(u),\varepsilon u)] = \hat{\partial}_L\left(\frac{1}{r}G_l^B(u)\right) + H_l^B(r,u), \tag{6.6}$$

where,

$$G_l^B(u) = \int_{-\infty}^{u} ds F(\varphi(s),\varepsilon s)(u-s)^{B-k+l+2}, \tag{6.7a}$$

$$H_l^B(r,u) = -\int_{-\infty}^{u} ds F(\varphi(s),\varepsilon s)\hat{\partial}_L\frac{(v-s)^{B-k+l+2}}{r}. \tag{6.7b}$$

The particular part of the solution H is well-defined at $B=0$, so the limit may be taken directly without any alteration associated with the finite part operator. The homogeneous part of the solution G will be defined by analytic continuation from the region in which it is well-defined to $B\to 0$ [384]. We examine now in more detail the solutions to these equations in the three sub-regions, the near Far Zone ($r\ll M/\varepsilon$), the intermediate Far Zone ($r\sim M/\varepsilon$), and the very Far Zone ($r\gg M/\varepsilon$). In each of the sub-regions, we will be able to make further simplifications to the Blanchet and Damour integrals to obtain more explicit expressions for the sourced fields. To make these arguments more direct, we will make use of a scaled coordinate $\tilde{r}=\varepsilon r$, for which the limits become $\tilde{r}\ll M$, $\tilde{r}\sim M$, and $\tilde{r}\gg M$, respectively.

In each of the sub-regions, the resulting nonlinear source behaves differently between the case of a quasistatic source and the case of an oscillatory source. We will therefore discuss the solutions

in each of the three sub-regions separately for the quasistatic source,

$$\bar{F} = \langle F(\varphi, \tilde{u}) \rangle \, (\varepsilon u), \tag{6.8}$$

and for the oscillatory source,

$$\tilde{F} = F - \bar{F} \equiv e^{\varphi(\tilde{u})/\varepsilon} f(\tilde{u}). \tag{6.9}$$

Notably, we find that the full retarded solution $\hat{\partial}_L(G_l(u)/r) + H_l(r,u)$ is required to ensure reasonable behavior of the approximation method in the near Far Zone for the quasistatic modes. However, in general, the inclusion or exclusion of the G-dependent part of the solution is a matter of choice: it is a homogeneous solution, and in principle is fixed in combination with an additional homogeneous solution by boundary conditions. Due to the resulting convenience of the calculational details, we choose to include the G-dependent part of the solution only for the quasistatic solution, and to solve for H alone for the oscillatory part of the solution. In the Post-Minkowski expansion, there is significantly different behavior for the perturbations sourced by rapidly varying modes, which tend to average over long integrals to small values, and quasistatic modes, which tend to accumulate over long scales, but have suppressed derivatives due to their long-scale dependence.

6.2.1 Simplification methods for quasistatic modes

The computation of the quasistatic modes deliberately includes both the homogeneous (6.7a) and the particular (6.7b) contributions to the sourced integral. In the calculation of the relevant integral,

$$\Box^{-1}_{\text{ret}}[r^{B-k}\hat{n}_L \bar{F}(\varepsilon u)] = \text{FP} \lim_{B\to 0} \frac{1}{K_k^B} \int_{-\infty}^{u} ds F(\varepsilon s) \hat{\partial}_L \frac{(u-s)^{B-k+l+2} - (v-s)^{B-k+l+2}}{r}, \tag{6.10}$$

we use a scaled integration variable $s \to \tilde{s} = \varepsilon s$, which more directly parameterizes the dependence of the source,

$$\Box^{-1}_{\text{ret}}[r^{B-k}\hat{n}_L \bar{F}(\varepsilon u)] = \text{FP} \lim_{B\to 0} \frac{1}{K_k^B} \int_{-\infty}^{\tilde{u}} d\tilde{s} \varepsilon^{k-l-3-B} F(\tilde{s}) \hat{\partial}_L \frac{(\tilde{u}-\tilde{s})^{B-k+l+2} - (\tilde{v}-\tilde{s})^{B-k+l+2}}{r}. \tag{6.11}$$

In the near Far Zone, $r \ll M/\varepsilon$, so $\tilde{r} \ll M$. The relative smallness of the scaled radial coordinate grants a near-cancellation between the retarded and advanced times which appear in the integral (6.11). The near-cancellation may then be expanded to obtain a reasonable value in the near Far Zone by methods closely similar to those used in [383]. The result is an appropriate suppression of

the resulting integral for higher modes in k, despite the apparent potentially high power of ε^{-1} in the integral (6.11).

In the very Far Zone, we once again rely on the scaled coordinate, except we wish to take the limit $r \gg M/\varepsilon$, or $\tilde{r} \gg M$, so we apply the scaled coordinate even to the radial factor and the dependence of the angular derivatives ∂_L. We then use the integral expression,

$$\Box^{-1}_{\rm ret}[r^{B-k}\hat{n}_L \bar{F}(\varepsilon u)] = {\rm FP}\lim_{B\to 0}\frac{1}{K^B_k}\int_{-\infty}^{\tilde{u}} d\tilde{s}\varepsilon^{k-2-B}F(\tilde{s})\hat{\tilde{\partial}}_L\frac{(\tilde{u}-\tilde{s})^{B-k+l+2}-(\tilde{v}-\tilde{s})^{B-k+l+2}}{\tilde{r}}. \quad (6.12)$$

Critically, as we are assuming a radial value large compared to the radiation reaction time, the integral gives an appropriate promotion of orders in the mass ratio ε from sources with larger k. The suppression in ε arises from the rapid falloff of sources with higher powers in r^{-1} on large distance scales.

6.2.2 Simplification methods for oscillatory modes

For the oscillatory modes, we explicitly compute only the particular solution $H_l(r,u)$ described in (6.7b), and absorb the homogeneous contribution to the metric perturbations $\hat{\partial}_L(G^B_l(u)/r)$ into the overall homogeneous perturbations (6.1) fixed by matching to the Interaction Zone result. Therefore, the single integral we wish to study is,

$$H_l(r,u) = -\int_{-\infty}^{u} ds e^{i\varphi(\varepsilon s)/\varepsilon}F(\varepsilon s)\frac{(v-s)^{l+2-k}}{r} \quad (6.13)$$

In many cases, the integrals of the form (6.13) can be significantly simplified by repeatedly integrating by parts, obtaining integrals of the factor $e^{i\varphi(\varepsilon s)/\varepsilon}$, and derivatives of the remaining factors. These tend to impose additional factors of ε/Ω. The careful expansion of each of the important orders of the Post-Minkowski expansion (6.2) will be presented in the publication [386]. Notably, the characteristics of the integrand are distinct in the very Far Zone limit $r \gg M/\varepsilon$ as compared to the near Far Zone limit $r \ll M/\varepsilon$, which has an impact on the functional dependence in the matching from the Interaction Zone and the extraction of the asymptotic perturbation.

6.2.3 Post-Minkowski summary

The results, which will be presented in the forthcoming publication [386] will give full details of the computation, including the necessary matching to the second order quasistatic Interaction Zone

modes and the promotion of hereditary effects in the very Far Zone. Under the appropriate treatment of the homogeneous modes separately for the oscillatory and the quasistatic modes, we obtain convergence in both the expansion in the mass ratio ε and in powers of the central mass M. Both forms of convergence are important, as they permit a terminating expansion of sources necessary to compute the metric perturbation to any precision in each of the sub-regions.

As a closing note to the description of the Post-Minkowski computation, we wish to emphasize that the results indicate that the rapidly varying waves are entirely fixed by matching to the interaction region, and in turn entirely fix the asymptotic values of the rapidly varying components. The results in this section then indicate that the only information that is transmitted from the Far Zone to the Interaction Zone is that information contained in the quasi-stationary solution at subleading order. These important feature of the dynamics far from the inspiral was the key motivation in developing the alternative description of the Far Zone in terms of a modified geometric optics formalism, described in the following section.

6.3 Multiscale geometric optics

The second method we have developed for the computation of the metric perturbations in the region $r \gg M$ is based closely on the geometric optics expansion, which is often used to approximate weak, high-frequency waves perturbing a background spacetime [387–392]. The basic form of our geometric optics expansion assumes a rapidly varying phase Θ, such that the scale of Θ is far less than the characteristic scale of the background curvature.

The technique of geometric optics for waves propagating on a background curvature of characteristic scale L far greater than the wavelength $\lambda \ll L$ is a well-documented method for computing the dynamics of rapidly varying perturbations [393,394]. At leading order, geometric optics calculations yield waves which travel along geodesics of the background, and further curvature corrections arise at subleading order in the ratio of lengthscales λ/L. In the particular case of a geometric optics expansion of a weak metric perturbation, we also obtain corrections from the nonlinear products of the leading wavelike perturbations acting as sources to the higher-order modes.

6.3.1 General formalism

In addition to the rapidly varying phase dependence, the metric perturbation in the Far Zone possesses a slow scale of variation, including variation in the retarded time on the scale of the radiation-reaction time M/ε. The propagating metric perturbations depend on a full set of scaled coordinates $\{\tilde{x}^\mu\} = \{\varepsilon x^i, \varepsilon u\}$ and a single rapidly varying phase variable Θ/ε, on which all physical variables depend 2π-periodically. We define a wave vector associated with the phase variable,

$$\nabla_\mu \Theta = l_\mu \tag{6.14}$$

Using the rescaled set of coordinates, we take the metric perturbation ansatz,

$$g_{\alpha\beta} = \varepsilon^{-2}\left(\eta_{\alpha\beta} + \varepsilon h_{\alpha\beta}(\tilde{x}^i) + \varepsilon^2 j_{\alpha\beta}\left(\tilde{x}^\mu, \frac{\Theta}{\varepsilon}\right) + \varepsilon^3 k_{\alpha\beta}\left(\tilde{x}^\mu, \frac{\Theta}{\varepsilon}\right) + \varepsilon^4 l_{\alpha\beta}\left(\tilde{x}^\mu, \frac{\Theta}{\varepsilon}\right) + \mathcal{O}(\varepsilon^5)\right). \tag{6.15}$$

In the ansatz (6.15), we have built in the physical criteria that a) the metric perturbations are purely outgoing (which is justified by the convergence of the perturbation theory), b) those outgoing waves diminish as at least r^{-1} (optionally with logarithmic dependence in r), and c) the leading perturbation on top of the Kerr background has scale $\sim \varepsilon/r = \varepsilon^2/\tilde{r}$, so appears in $j_{\alpha\beta}$. The contribution $h_{\alpha\beta}$ is entirely fixed by an expansion of the Kerr background at large radii, while the $j_{\alpha\beta}$ and $k_{\alpha\beta}$ contain both contributions from the expansion of the static Kerr background and dynamical perturbations determined by matching to the interaction region and expansion of the wave equation in the geometric optics approximation. The remaining general form of the expansion is given in Section 4.5. In this section, we present the remaining higher order expansion of the geometric optics approximation necessary for post-adiabatic computations.

6.3.2 First subleading order corrections

Recall that at first subleading order, the expansion of the Einstein field equation expands to,

$$\delta G^{(-2)}_{\alpha\beta}[k] + \delta G_{\alpha\beta}[h] + \delta G^{(-1)}_{\alpha\beta}[j] + 2\delta^2 G^{(-2)}_{\alpha\beta}[h, j] = 0. \tag{6.16}$$

In Section 4.5.8, we used two components of this equation to determine the radial dependence of the metric perturbations j_{mm} and $j_{\bar{m}\bar{m}}$,

$$\delta G^{(-2)}_{mm}[k] + \delta G^{(-1)}_{mm}[j] + 2\delta^2 G^{(-2)}_{mm}[h,j] = \frac{M}{\tilde{r}}\delta j''_{mm} - \frac{1}{2\tilde{r}}\delta j'_{mm} + \frac{1}{2}\partial_{\tilde{v}} j'_{mm} = 0, \tag{6.17a}$$

$$\delta G^{(-2)}_{\bar{m}\bar{m}}[k] + \delta G^{(-1)}_{\bar{m}\bar{m}}[j] + 2\delta^2 G^{(-2)}_{\bar{m}\bar{m}}[h,j] = \frac{M}{\tilde{r}}\delta j''_{\bar{m}\bar{m}} - \frac{1}{2\tilde{r}}\delta j'_{\bar{m}\bar{m}} - \frac{1}{2}\partial_{\tilde{v}} j'_{\bar{m}\bar{m}} = 0, \tag{6.17b}$$

which solved to yeild a mode decomposition, $j^{\ell m}_{mm}(\Theta,\tilde{r},\tilde{v}) = \sum_k e^{ik\Omega(\tilde{u})\Theta} R_k(\tilde{r},\tilde{u})$, where the radial function was solved using (6.17) giving

$$R_k(\tilde{r},\tilde{u}) = \frac{R(r_0,\tilde{u}) r_0}{\tilde{r}} \tilde{r}^{2ik\Omega(\tilde{u})M}. \tag{6.18}$$

The final three components of the first subleading order geometric optics equations involve both the rapidly varying components of δk and those of δj,

$$-\frac{1}{6}k''_{m\bar{m}} - \frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{nm} - \frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{n\bar{m}} + \frac{j'_{nn}}{2\tilde{r}} - \frac{1}{2}\bar{m}^\mu \partial_\mu j'_{nm} - \frac{1}{2} m^\mu \partial_\mu j'_{n\bar{m}} = 0 \tag{6.19a}$$

$$-\frac{1}{12}k''_{lm} - \frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{mm} + \frac{j'_{nm}}{4\tilde{r}} - \frac{1}{4}\bar{m}^\mu \partial_\mu j'_{mm} - \frac{1}{4} l^\mu \partial_\mu j'_{nm} = 0, \tag{6.19b}$$

$$-\frac{1}{12}k''_{l\bar{m}} - \frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{\bar{m}\bar{m}} + \frac{j'_{n\bar{m}}}{4\tilde{r}} - \frac{1}{4} m^\mu \partial_\mu j'_{\bar{m}\bar{m}} - \frac{1}{4} l^\mu \partial_\mu j'_{n\bar{m}} = 0, \tag{6.19c}$$

Several of the δk components can be removed by fixing the remaining gauge freedom in δj. We take advantage of our remaining gauge freedom, which allows us to arbitrarily set each component of $\delta j_{n\mu}$, to entirely eliminate the portions of the equations (6.19) which depend on δj. Therefore, our remaining three gauge conditions for δj are,

$$-\frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{nm} - \frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{n\bar{m}} + \frac{j'_{nn}}{2\tilde{r}} - \frac{1}{2}\bar{m}^\mu \partial_\mu j'_{nm} - \frac{1}{2} m^\mu \partial_\mu j'_{n\bar{m}} = 0 \tag{6.20a}$$

$$-\frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{mm} + \frac{j'_{nm}}{4\tilde{r}} - \frac{1}{4}\bar{m}^\mu \partial_\mu j'_{mm} - \frac{1}{4} l^\mu \partial_\mu j'_{nm} = 0, \tag{6.20b}$$

$$-\frac{\cot(\theta)}{2\sqrt{2}\tilde{r}} j'_{\bar{m}\bar{m}} + \frac{j'_{n\bar{m}}}{4\tilde{r}} - \frac{1}{4} m^\mu \partial_\mu j'_{\bar{m}\bar{m}} - \frac{1}{4} l^\mu \partial_\mu j'_{n\bar{m}} = 0, \tag{6.20c}$$

These three equations fix the remaining components δj_{nn}, δj_{nm}, and $\delta j_{n\bar{m}}$ left undetermined by the gauge choice at the previous order $\delta j_{nl} = 0$ discussed in Section 4.5.

Combining, then the three equations (6.19) with the final remaining component of the Einstein field equations, we obtain the tidy result $l^\nu \delta k_{\nu\mu} = 0$, $\eta^{\mu\nu}\delta k_{\mu\nu} = 0$, just as we had for the leading order perturbations δj. At this point in the computation, we still have the gauge freedom in the

$n^\mu \delta k_{\mu\nu}$ components of the metric perturbation. If we wish to push the expansion to higher order, we would once again set $\delta k_{nl} = 0$ to ensure consistency with the Lorenz gauge and reserve the remaining gauge freedom to simplify the equations for higher order metric perturbations. However, since the first subdominant order of outgoing waves is all that is required for a post-adiabatic waveform, for this work we immediately take the simpler condition that $n^\mu \delta k_{\mu\nu} = 0$. The choice $n^\mu \delta k_{\mu\nu}$ leaves only the δk_{mm} and $\delta k_{\bar{m}\bar{m}}$ modes to propagate to future null infinity. As in the leading order rapidly varying modes δj, the subleading modes δk are determined by matching to the Interaction Zone, and the radial dependence will be extracted from equations analogous to (4.114), but at the next order in the expansion of small mass ratio.

6.3.3 Extracting the required second subleading order information

Some of the information in the second subleading order geometric optics wave equation will be important to a computation of post-adiabatic effects. First, we still have not yet determined the quasistatic part of the leading nontrivial metric perturbation, $j_0(\tilde{x}^\mu)$, which is only determined by the second subleading order of wave equation. In addition, however, it is important to determine any correction to the radial propagation of second order modes $\delta k_{\bar{m}\bar{m}}$. Consider first the general form of the second subleading order wave equation,

$$\begin{aligned} 0 =& \delta G^{(0)}_{\mu\nu}[j] + \delta G^{(-1)}_{\mu\nu}[k] + \delta G^{(-2)}_{\mu\nu}[l] + 2\delta^2 G^{(-1)}_{\mu\nu}[h, j] \\ &+ 2\delta^2 G^{(-2)}_{\mu\nu}[h, k] + \delta^2 G^{(0)}_{\mu\nu}[h, h] + \delta^2 G^{(-2)}_{\mu\nu}[j, j] + 3\delta^3 G^{(-2)}_{\mu\nu}[h, h, j] \end{aligned} \tag{6.21}$$

The first step is to determine the j_0 modes by taking an average over Θ, giving the simplified equation,

$$\delta G^{(0)}_{\mu\nu}[j_0] + \delta^2 G^{(0)}_{\mu\nu}[h, h] + \left\langle G^{(-2)}_{\mu\nu}[\delta j, \delta j] \right\rangle = 0 \tag{6.22}$$

We now make the further split of the quasistatic modes of the second order metric perturbation into those determined by the background Kerr metric j_0^B, and the remaining part to be determined by the wave equation j_0^R,

$$j_0 = j_0^B + j_0^R. \tag{6.23}$$

Using this split, we can immediately take advantage of the satisfaction of the perturbative Einstein field equation by the full perturbatively expanded Kerr solution, giving

$$\delta G^{(0)}_{\mu\nu}[j_0^B] + \delta^2 G^{(0)}_{\mu\nu}[h,h] = 0. \tag{6.24}$$

Using (6.24) in (6.22), we obtain the simplified equation for the quasistatic modes which will be used in the matching back to the Interaction region,

$$\delta G^{(0)}_{\mu\nu}[j_0^R] + \left\langle G^{(-2)}_{\mu\nu}[\delta j, \delta j] \right\rangle = 0. \tag{6.25}$$

We note that in the Far Zone and in the Lorenz gauge which we impose for j_0 (see Section 4.5.5), the leading wave operator $\delta G^{(0)}_{\mu\nu}$ may be expressed as the simple wave operator, giving the differential equation,

$$\Box j^R_{0\,\mu\nu} = -\left\langle \delta^2 G^{(-2)}_{\mu\nu}[\delta j, \delta j] \right\rangle. \tag{6.26}$$

The wave equation (6.26) may be solved using standard Post-Minkowski techniques, such as those presented in [384] and reviewed in Section 6.2 for precisely this type of equation and limit. In particular, the solution may be expressed as the integral,

$$j_0^R = \mathrm{FP}\lim_{B\to 0} \frac{1}{K_B} \int_{-\infty}^{\tilde{u}} d\tilde{s} \left\langle \delta^2 G^{(-2)}_{\mu\nu}[\delta j, \delta j] \right\rangle \tilde{\partial}_L \frac{(\tilde{u}-\tilde{s})^{B-k+l+2} - (\tilde{v}-\tilde{s})^{B-k+l+2}}{\tilde{r}}. \tag{6.27}$$

Evaluating the quasistatic integral (6.27) in the small $\tilde{r}$ limit, and replacing $\tilde{r} \to \varepsilon r$ gives a near-cancellation between the $\tilde{u}$ and $\tilde{v}$ contributions to the integral, ensuring that the dominant result is suppressed by at least one order in the mass ratio ε. This is as we should anticipate; the quadratic combination of the leading oscillatory modes is a second-order perturbation in the Interaction Zone, so an appropriate matching of these additional quasistatic modes is then possible in the Interaction Zone limit. It is notable, however, that the integral (6.27) offers no such suppression for evaluation at large $\tilde{r}$

The integral (6.27) determines the final piece of information for the full second order metric perturbation j, which then provides important information to the Interaction Zone in the quasistatic modes, and determines the asymptotic behavior of the modes for waveform extraction at future null infinity. The additional quasistatic modes determined by (6.27) are a hereditary effect: they slowly accumulate over the duration of the inspiral, and can have nonzero value even at retarded times significantly greater than the termination of the inspiral.

The final contributions to the wave equations which we will consider at second subleading order are those which determine the radial dependence of the periodic perturbations δk. All other contributions not yet discussed are of such an order that they will not contribute to post-adiabatic computations of the waveform or dynamical gauge invariants, so are disregarded. Taking the $m\,m$ and $\bar{m}\,\bar{m}$ components of the wave equation (6.21), we obtain, respectively,

$$\begin{aligned}0 =&\frac{M}{3r}\delta k''_{mm} - \frac{1}{6r}\delta k'_{mm} - \frac{1}{6}l^\alpha\nabla_\alpha\delta k'_{mm} + \frac{M}{2r^2}\delta j'_{mm} - \frac{\sqrt{2}M\cot(\theta)}{r^2}\delta j'_{nm}\\ &- \frac{2M}{r}n^\alpha\nabla_\alpha\delta j'_{mm} + \frac{2M}{r}m^\alpha\nabla_\alpha\delta j'_{nm} - \frac{1}{2r^2}\delta j_{mm} - \frac{\cot^2(\theta)}{2r^2}\delta j_{mm} + \frac{\csc^2(\theta)}{2r^2}\delta j_{mm}\\ &+ \frac{\cot(\theta)}{2\sqrt{2}r^2}\delta j_{nm} - \frac{1}{4r}l^\alpha\nabla_\alpha\delta j_{mm} + \frac{1}{2r}n^\alpha\nabla_\alpha\delta j_{mm} + \frac{\cot(\theta)}{2\sqrt{2}r}l^\alpha\nabla_\alpha\delta j_{nm} - \frac{1}{r}m^\alpha\nabla_\alpha\delta j_{nm}\\ &- \frac{1}{4}\nabla_\alpha\nabla^\alpha\delta j_{mm} + \frac{1}{2}m^\alpha\bar{m}^\beta\nabla_\beta\nabla_\alpha\delta j_{mm} - \frac{1}{2}l^\alpha m^\beta\nabla_\beta\nabla_\alpha\delta j_{nm}\end{aligned} \tag{6.28a}$$

$$\begin{aligned}0 =&\frac{M}{3r}\delta k''_{\bar{m}\bar{m}} - \frac{1}{6r}\delta k'_{\bar{m}\bar{m}} - \frac{1}{6}l^\alpha\nabla_\alpha\delta k'_{\bar{m}\bar{m}} + \frac{M}{2r^2}\delta j'_{\bar{m}\bar{m}} - \frac{\sqrt{2}M\cot(\theta)}{r^2}\delta j'_{n\bar{m}}\\ &- \frac{2M}{r}n^\alpha\nabla_\alpha\delta j'_{\bar{m}\bar{m}} + \frac{2M}{r}\bar{m}^\alpha\nabla_\alpha\delta j'_{n\bar{m}} - \frac{1}{2r^2}\delta j_{\bar{m}\bar{m}} - \frac{\cot^2(\theta)}{2r^2}\delta j_{\bar{m}\bar{m}} + \frac{\csc^2(\theta)}{2r^2}\delta j_{\bar{m}\bar{m}}\\ &+ \frac{\cot(\theta)}{2\sqrt{2}r^2}\delta j_{n\bar{m}} - \frac{1}{4r}l^\alpha\nabla_\alpha\delta j_{\bar{m}\bar{m}} + \frac{1}{2r}n^\alpha\nabla_\alpha\delta j_{\bar{m}\bar{m}} + \frac{\cot(\theta)}{2\sqrt{2}r}l^\alpha\nabla_\alpha\delta j_{n\bar{m}} - \frac{1}{r}\bar{m}^\alpha\nabla_\alpha\delta j_{n\bar{m}}\\ &- \frac{1}{4}\nabla_\alpha\nabla^\alpha\delta j_{\bar{m}\bar{m}} + \frac{1}{2}\bar{m}^\alpha m^\beta\nabla_\beta\nabla_\alpha\delta j_{\bar{m}\bar{m}} - \frac{1}{2}l^\alpha\bar{m}^\beta\nabla_\beta\nabla_\alpha\delta j_{n\bar{m}}\end{aligned} \tag{6.28b}$$

The equations (6.28) then determine the remaining unfixed coordinate dependence of the rapidly varying second order modes δk, up to boundary conditions determined by matching with the Interaction Zone.

In the forthcoming work [386], we further develop the geometric optics formalism, explore the connections between this formalism and the closely related Post-Minkowski expansion discussed in the previous section, and determine alterations required for extracting an asymptotically flat solution in the limit $\tilde{r} \to \infty$. Furthermore, we determine the matching conditions to the Interaction Zone necessary to supply the information from the quasistatic modes to the multiscale approximation and extract the information from the oscillatory modes in the Interaction Zone to the Far Zone.

6.4 | Conclusions

In various chapters of this dissertation, I have presented a much advanced picture of the multiscale approximation techniques for extreme mass ratio inspirals. In Chapter 4, we have laid out the

foundations necessary for an Adiabatic order computation using multiscale techniques, In Chapter 5, we have presented the full computation details necessary to develop a multiscale computation in the Interaction region, away from the Far and Near Horizon Zones. Finally, in this chapter, I have presented the bulk of the computation necessary to evaluate the propagating modes in the Far Zone.

A full development of the multiscale approximation method is of great importance for highly accurate simulation of extreme mass ratio inspirals, as it is the only currently proposed method that captures all of the required effects for a post-adiabatic waveform. In addition, it holds the significant promise of offering methods for improving computational efficiency, by taking full advantage of the perturbative nature of the separation of timescales. In particular, multiscale approximation methods are amenable to the near-identity techniques recently showed to offer vast performance improvements [395], by a factor comparable to ε^{-1}.

To close my discussion of the multiscale approximation methods, developed over the last three chapters, I summarize our current understanding of the algorithm which should be followed to compute the multiscale post-adiabatic waveform. For each step of the computation, I either make reference to the relevant derivation presented in this dissertation, or note the future work which will address the computation in greater detail. Due to the bias of current methods, I explain the waveform generation process assuming that Teukolsky variables will be used throughout, so reconstruction steps are required. The computation of dynamical invariants through second order in the Interaction Zone should likely follow a similar flow, although it is possible that such computations will prefer a Lorenz gauge wave equation, as there is currently no known formulation of second order metric reconstruction.

Post-adiabatic waveforms

Offline step: Generate parameter space of periodic orbits and perturbations for each parameter set in a sufficiently dense sample of the space, $\{\varepsilon, a, \delta M, \delta a, J'^{M(0)}, J'^{M(1)}\}$ **(where** P'^{M} **are determined via near-identity transformation, see Section 5.A)**

1. (Interaction Zone) Use leading parameters $\{\varepsilon, a, P'^{M(0)}\}$ to determine the leading perturbations $\psi_0^{(1)}$ and $\psi_4^{(1)}$ via the Teukolsky equation with a pointlike source: Section 5.3 for wave equation, implementation to follow standard self force methods [396].

2. (Interaction Zone) Use leading Teukolsky variables $\psi_0^{(1)}$ and $\psi_4^{(1)}$ to reconstruct the leading metric perturbation $h_{\mu\nu}^{(1)}$ via reconstruction algorithm : Section 5.3 (originally developed in publications [397, 398]).

3. (Multiscale Orbit) Use first order metric perturbations $h_{\mu\nu}^{(1)}$ to determine the leading slow evolution of near-identity transformed $P'^{M(0)}$: Section 5.A for near-identity details, Section 4.4 for a discussion of the self-force evolution of the orbit derived in [399].

4. (Multiscale Orbit) Use first order metric perturbations $h_{\mu\nu}^{(1)}$ and $\{J'^{M(1)}\}$ to determine the $\mathcal{O}(\varepsilon)$ frequency corrections from self force effects, g^A, and the $\mathcal{O}(\varepsilon)$ frequency corrections from $\omega^A(J^M)$: Sections 4.4 and 5.A.

5. (Interaction Zone) Use first order metric perturbations $h_{\mu\nu}^{(1)}$ to determine the slow evolution of the central mass δM and spin δa: Section 5.3.5

6. (Interaction Zone) Use subleading frequency values $\Omega^{A(1)}$, parameters $J'^{M(1)}$, and self-force values $g^{(1)A}$ and $G^{(1)M}$ to determine the worldline correction $z^{(1)\mu}(\tilde{w}, q^A)$, at fixed slow time $\tilde{w}$: Section 5.A.

7. (Interaction Zone) Use the subleading worldline value $z^{(1)}$ and the regular metric perturbation $h_{\mu\nu}^{\mathcal{R}(1)}$ to determine the subleading puncture $h_{\mu\nu}^{\mathcal{P}(2)}$: forthcoming publication [400].

8. (Interaction Zone) Use $h_{\mu\nu}^{(1)}$ to determine the multiscale-adjusted effective Teukolsky-Lousto-Campanelli source: Section 5.3

9. (Near-Horizon Zone) Determine Near-horizon amplitudes from $h_{\mu\nu}^{(1)}$ derived in the Interaction Zone: future publication [401]

10. (Near-Horizon Zone) Determine any quasistatic, sourced second-order perturbations $h_{\mu\nu}^{(2)}$ from nonlinear sources in the Near-Horizon limit: future publication [401]

11. (Far Zone) Determine Far Zone amplitudes from $h^{(1)}_{\mu\nu}$ derived in the Interaction Zone: Section 4.5 and further matching details in forthcoming publication [386]

12. (Far Zone) Determine quasistatic, sourced second-order perturbations $h^{(2)}_{\mu\nu}$ from nonlinear sources in the Far Zone: Section 6.3.3

13. (Interaction Zone) Use the method of extended particular solutions and boundary conditions supplied by quasistatic Near Horizon and Far Zone computations to determine the second order Teukolsky amplitudes $\psi^{(2)}_0$, $\psi^{(2)}_4$: Partial details in Section 5.3.4, full exploration in future work.

14. (Interaction Zone, Near-Horizon Zone, Far Zone) Use Teukolsky-Lousto-Campanelli quantities and first order quantities to evaluate the integrals needed to determine the slow evolution of post-adiabatic near-identity transformed orbit parameters $P'^{M(1)}$: Section 5.4, further details in forthcoming publication [400]

15. (Far Zone) Match second order solutions $\psi^{(2)}_0$, $\psi^{(2)}_4$ to Far Zone perturbations δk, and extract asymptotic waveform at fixed $\tilde{w}$: δk radial dependence from Section 6.3.3, matching details in forthcoming publication [386]

Offline output: Slow evolutions $\partial_{\tilde{w}}\delta a, \partial_{\tilde{w}}\delta M$ (Step 5), $\partial_{\tilde{w}}P'^{M(0)}$ (Step 3), $\partial_{\tilde{w}}P'^{M(1)}$ (Step 14), leading frequencies ω^A, subleading frequencies $\Omega^{A(1)}(\tilde{w})$ (Step 4), fixed-$\tilde{w}$ waveform through second order $\delta j(q^A, \tilde{w})$, $\delta k(q^A, \tilde{w})$ (Steps 11,15)

Online step: given a particular starting point in the parameter space $\{\varepsilon, a, \delta M(\tilde{w}_0), \delta a(\tilde{w}_0), P'^{M(0)}(\tilde{w}_0), P'^{M(1)}(\tilde{w}_0)\}$, compute a full inspiral

1. For each slow time, lookup the values of the instantaneous evolutions of the relevant parameters in the generated parameter space from the offline step, and evolve all values in slow time $\tilde{w} \to \tilde{w} + \Delta\tilde{w}$, generating a full inspiral of slow-time dependent values $\{\delta M(\tilde{w}), \delta a(\tilde{w}), P'^{M(0)}(\tilde{w}), P'^{M(1)}(\tilde{w}), \Omega^{A(0)}(\tilde{w}), \Omega^{A(1)}(\tilde{w})\}$

2. Integrate $dq^A/dw = \Omega^{A(0)}(\tilde{w}) + \varepsilon\Omega^{A(1)}(\tilde{w})$ and $\tilde{w} = \varepsilon w$ to determine $q^A(w)$ over the full inspiral

3. Extract the asymptotic waveform by evaluating $\delta j(q^A(w), \varepsilon w) + \varepsilon\delta k(q^A(w), \varepsilon w)$

Online output: A full waveform with errors of $\mathcal{O}(\varepsilon^2)$ in amplitude and $\mathcal{O}(\varepsilon)$ in phase

Chapter 6 Bibliography

[383] A. Pound, *Second-order perturbation theory: problems on large scales*, Phys. Rev. **D92** (2015) (10), p. 104047

[384] L. Blanchet and T. Damour, *Hereditary effects in gravitational radiation*, Phys. Rev. **D46** (1992), pp. 4304–4319

[385] L. Blanchet and T. Damour, *Tail Transported Temporal Correlations in the Dynamics of a Gravitating System*, Phys. Rev. **D37** (1988), p. 1410

[386] E. E. Flanagan, T. Hinderer, J. Moxon and A. Pound, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. III. Dynamics of the far zone waves* ((In prep.))

[387] R. A. Isaacson, *Gravitational Radiation in the Limit of High Frequency. I. The Linear Approximation and Geometrical Optics*, Phys. Rev. **166** (1967), pp. 1263–1271

[388] F. A. Handler and R. A. Matzner, *GRAVITATIONAL WAVE SCATTERING*, Phys. Rev. **D22** (1980), pp. 2331–2348

[389] C. Bloomer, *Optical Geometry of the Kerr Space-time* (2011)

[390] S. R. Dolan, *Higher-order geometrical optics for electromagnetic waves on a curved spacetime* (2018)

[391] S. R. Dolan, *Geometrical optics for scalar, electromagnetic and gravitational waves in curved spacetime*, in *Int. J. Mod. Phys. D 1843010 (2018), Int. J. Mod. Phys. D 1843010 (2018)* (2018)

[392] Z. Chang, C.-G. Huang and Z.-C. Zhao, *Motion of photons in a gravitational wave background*, Chin. Phys. **C41** (2017) (9), p. 093108

[393] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973)

[394] R. M. Wald, *General Relativity* (Chicago Univ. Pr., Chicago, USA, 1984)

[395] M. Van De Meent and N. Warburton, *Fast Self-forced Inspirals* (2018)

[396] B. Wardell, *Self-force: Computational Strategies*, Fund. Theor. Phys. **179** (2015), pp. 487–522

[397] C. O. Lousto and B. F. Whiting, *Reconstruction of black hole metric perturbations from Weyl curvature*, Phys. Rev. **D66** (2002), p. 024026

[398] A. Ori, *Reconstruction of inhomogeneous metric perturbations and electromagnetic four potential in Kerr space-time*, Phys. Rev. **D67** (2003), p. 124010

[399] T. Hinderer and E. E. Flanagan, *Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime: Orbital motion*, Phys. Rev. D **78** (2008), p. 064028,
URL: http://link.aps.org/doi/10.1103/PhysRevD.78.064028

[400] E. E. Flanagan, T. Hinderer, J. Moxon and A. Pound, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. II. Dynamics of the strong-field region* ((In prep.))

[401] S. Isoyama, A. Pound, T. Tanaka and K. Yamada, *The two body problem in general relativity in the extreme mass ratio limit via multiscale expansions. IV. Dynamics of the near-horizon region* ((In prep.))

Chapter 7

Conjugate constraints and modified Dirac brackets

Coauthor:

Éanna Flanagan, Cornell University

7.1 Introduction

The Hamiltonian formulation for the dynamics of mechanical and field systems is a valuable complementary alternative to the Lagrangian formalism. For the method of canonical quantization, the Hamiltonian formulation is more desirable, due to the ready definitions of the Poisson brackets which are often directly promoted to canonical commutation relations,

$$\{p^\mu, q^\nu\} = \delta^{\mu\nu} \rightarrow [p^\mu, q^\mu] = -i\hbar\delta^{\mu\nu}. \tag{7.1}$$

In the Lagrangian formalism, for a finite-dimensional system, the independent variables are the set of N position variables q^μ and their time derivatives $\dot{q}^\mu$, and the equations of motion are the Euler-Lagrange equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\mu}\right) - \frac{\partial L}{\partial q^\mu} = 0, \tag{7.2}$$

which are often second-order differential equations. By contrast, the Hamiltonian formalism derives the dynamics on a phase space of $2N$ variables q^μ and their conjugate momenta p^μ. The result is

the Hamilton equations of motion:

$$\dot{q}^\mu = \frac{\partial H}{\partial p_\mu}, \tag{7.3a}$$

$$\dot{p}^\mu = -\frac{\partial H}{\partial q^\mu}, \tag{7.3b}$$

which typically gives a description in terms of $2N$ first-order differential equations.

The conversion from the Lagrangian to the Hamiltonian description proceeds gracefully unless there is a subset of the Euler-Lagrange equations (7.2) which cannot be represented as a set of Hamiltonian equations. One of the clearest immediate examples of such an equation is Gauss' Law from electrodynamics,

$$\partial_i E^i = \frac{\rho}{4\pi\varepsilon_0}, \tag{7.4}$$

which arises cleanly from the Euler-Lagrange equations (7.2) for electrodynamics. However, as the electric field E^i is the momentum conjugate to the spatial components of the vector potential A^i, Gauss' Law cannot be represented as any portion of the Hamilton equations of motion (7.3). Instead, it arises as an additional input generated by the conversion from the Lagrangian to the Hamiltonian formalism. Such additional restrictions are referred to in the Hamiltonian formalism as *constraints*. As with Gauss' law, the constraints of a Hamiltonian system often correspond to initial conditions for the dynamics.

Constraints arise as a result of the naive $2N$-dimensional phase space overcounting the number of true dynamical degrees of freedom encoded in the Lagrangian system. As a result, some of the spare degrees of freedom give rise to constraints, which then limit the phase space. There is therefore a deep connection between those systems which give rise to constraints in the Hamiltonian formalism and those systems which exhibit gauge symmetry. However, in some cases, the set of constraints possesses Poisson brackets which fail to be consistent in the conversion of those Poisson brackets to canonical commutation relations (7.1).

Multiple methods have been developed for working with gauge systems such that a quantized version of the theory becomes accessible. One well-studied method is Dirac quantization [402, 403], which is the main focus of this chapter. BRST quantization [404, 405], by contrast, introduces additional fictitious degrees of freedom to the system which have the effect of precisely canceling the unphysical degrees of freedom generated by a naive phase space construction. Due to the

availability of a generally covariant formulation for BRST quantization, it is often the preferred method for exploring quantum field theories. Alternatively, it is possible to construct the symplectic product on the constraint surface directly without first constructing the naive phase space, as is developed in [406]. This more direct approach may prove more difficult for theories in which only perturbative derivations are readily available, but the derivation of [406] offers fundamental insight to the general nature of Hamiltonian systems.

While significantly less popular in textbook treatments of quantum field theory, Dirac quantization techniques have been constructed for many field theories. In particular, the quantization procedure for canonical quantum gravity [407,408] shares many features with the Dirac quantization procedure, although with significant additional formal development. Additionally, perturbative field theories have also been treated with Dirac quantization [409–414].

Due to the benefit of having alternative descriptions of a system, and the ability of Dirac quantization to easily describe phenomena such as state dressing (see Section 7.4), we find it valuable to further explore Dirac quantization. We are particularly interested in the context of theories which generate nontrivial constraints, with the motivation that gauge theories can be expressed in that context. Due to the fundamental difficulties in executing the desired canonical quantization in certain systems, or in certain choices of physically equivalent gauges, Dirac has developed [402] a method of defining an appropriate refinement of the naive Poisson bracket product on phase space. The refined bracket, called the *Dirac bracket*, is constructed to reproduce the intrinsic bracket associated with the constraint surface (see Section **??**). However, it is carefully defined in directions off the constraint surface so that the various constraints then have vanishing brackets with one another, allowing the theory to be consistently quantized. In Section 7.2, we review the formalism of Hamiltonian constraints in general and the construction of Dirac brackets.

We present an alternative construction for deriving similar quantities in second-class systems and tools for working with first-class systems. Our construction of modified Dirac brackets is presented in Section 7.3. Notably, the modified Dirac bracket can also be used in the case of first-class constraints to convert a weakly vanishing Poisson bracket to a strongly vanishing modified Dirac bracket. In Section 7.4, we present a perturbative expansion for constructing a physical, constraint-satisfying state for the first-class case.

7.2 Constraints in Hamiltonian systems

7.2.1 Hamiltonian constraint derivation and classification

In this section, we discuss the general procedure for converting from a Lagrangian to a Hamiltonian description of a physical system, and in the process categorize the various constraints which can arise from that conversion. This overview follows closely the discussion of constraint systems in [403]. Consider a general Lagrangian $L(\dot{q}^\mu, q^\mu)$, for which the dynamics of the physical system are described by extremization of the action

$$S = \int dt L(\dot{q}^\mu, q^\mu). \tag{7.5}$$

Define, then, a set of canonical momenta,

$$p_\mu = \frac{\partial L}{\partial \dot{q}^\mu}. \tag{7.6}$$

Here, we will make use of greek indices to run over the full set of coordinates in the system, and we use Einstein summation convention for the indices. If the set of equations (7.6) gives rise to momenta which are all independent functions of coordinates and their derivatives, the system has no constraints, and the Hamilton equations of motion (7.3) entirely describe the evolution of the system. Consider the alternative case in which there are a set of relations

$$\chi^1_M(q,p) = 0, \tag{7.7}$$

for $M = 1, 2, \ldots, N_{\text{primary}}$, which are determined by the momentum definitions (7.6). The full, irreducible set of relations extracted directly from the definitions of canonical momenta (7.7) are called *primary constraints*. In simple examples, there are variables q^i in the set q^μ, such that $\dot{q}^i$ does not appear at all in the Lagrangian L. In such examples, one of the relations (7.7) is the vanishing of the corresponding canonical momentum $p^i = 0$.

With the canonical momenta determined by (7.6), we may proceed to define the Hamiltonian associated with this system. For the set of independent dynamical momentum degrees of freedom p_i, one can invert the set of equations (7.6) to determine $\dot{q}^i(p,q)$. The Hamiltonian is obtained via the Legendre transform,

$$H(p^\mu, q^\mu) = p_i \dot{q}^i(p,q) - L(\dot{q}^\nu(p^\mu, q^\mu), q^\nu) + C^M \chi^1_M(q,p), \tag{7.8}$$

where the coefficients C^M are included for full generality of the Hamiltonian. The Hamiltonian which includes the contributions C^M is often referred to as the 'extended' Hamiltonian. On the physical constraint surface described by (7.7), the time dependence described by the Hamiltonian (7.8) is unchanged by the inclusion of arbitrary contributions proportional to $\chi^1_M(q,p)$.

However, in general it is not sufficient to simply impose the relations (7.7) to define the primary constraint surface and to proceed to evaluate the dynamics of the system. We must also ensure that as the system evolves, it does not subsequently depart from the constraint surface. Formally, we must require that the time derivatives of the primary constraints vanish,

$$\dot{\chi}^1_M(q,p) = \{\chi^1_M, H\} = 0, \tag{7.9}$$

where the operation $\{\}$ is the Poisson bracket, defined as

$$\{F,G\} = \frac{\partial F}{\partial q^\mu}\frac{\partial G}{\partial p_\mu} - \frac{\partial F}{\partial p^\mu}\frac{\partial G}{\partial q_\mu}. \tag{7.10}$$

The set of independent conditions obtained from the first derivatives (7.9) of the primary constraints which are also independent of the primary constraints (7.7) are referred to as *secondary constraints* χ^2_N. The procedure of repeatedly taking the time derivatives of constraints to impose adhesion to the constraint surface is then iterated until no further independent conditions are generated, yielding *tertiary constraints* and higher order constraints. Note that at each step, we may either obtain new constraints or develop restrictions on the arbitrary coefficients C^M, and at each step the new constraints must be incorporated into the extended Hamiltonian for subsequent steps. Finally, the full set of constraints is obtained,

$$\chi_A = (\chi^1_M, \chi^2_N, \chi^3_O, \dots). \tag{7.11}$$

As a final notational point, it is useful to denote separately the equality of physical quantities in the full phase space, and equality only under the restriction to the constraint surface $\chi_A = 0$. The former is referred to as *strong equality*, and is denoted with an ordinary $=$. The latter is referred to as *weak equality*, and is denoted by $\approx$.

If the Poisson brackets of the various constraints with one another vanishes on the constraint surface,

$$\{\chi_A, \chi_B\} \approx 0, \tag{7.12}$$

the set of constraints are said to be *first-class*. If the matrix of the Poisson brackets of constraints is nondegenerate ($\mathrm{Det}[\chi_A, \chi_B] \neq 0$), the set of constraints is said to be *second-class*. If neither condition is true, then the set of constraints is said to be mixed, and there are transformations of the constraint system which yeild a subset of first-class and a subset of second-class constraints. These terms first and second class are easy to confuse with primary, secondary, etc., , but we'd like to emphasize that the distinction between first vs. second-class constraints has only to do with their structure associated with Poisson brackets, and has nothing to do with whether they are primary, secondary, etc. Because the first and second-class classification has to do with the structure of the dynamics, and primary, secondary, etc., has only to do with the derivation procedure, the distinction between primary, secondary, etc. is a less fundamental distinction.

7.2.2 Dirac quantization of a constraint systems

In this section, we review the Dirac methods of quantizing Hamiltonian systems [402] that exhibit first or second-class constraints introduced in the previous section. The methods depend significantly on whether the constraints are first-class or second-class, and second-class constraints require the additional definition of Dirac brackets, which become the foundation for the canonical commutation relations rather than the Poisson bracket. To more clearly understand the necessity of this redefinition, consider the toy example in which the pair of constraints is $q_1 = 0$ and $p_1 = 0$. We may not, then, simultaneously apply these definitions at operator level and enforce the desired canonical commutation relation,

$$[q_1, p_1] = i\hbar \tag{7.13}$$

First-class constraints, which have only weakly vanishing brackets (7.12), can be quantized without the introduction of additional structure to the commutation relations. When the set of Poisson brackets is promoted to commutation relations, a weakly vanishing set of brackets implies,

$$[\chi_A, \chi_B]\big|_{\chi_C=0} = 0. \tag{7.14}$$

Therefore, the constraint surface is well-defined in terms of the operators to which we promote the physical variables. Note that we don't necessarily want to impose the constraints at operator level, as one of the key features of a quantized theory is the possibility off-shell departures from strict

satisfaction of the equations of motion, but the physical states are required to remain on-shell. To maintain adherence to the constraint surface, Dirac has showed [402] that in the first-class case, it suffices to impose the set of constraints, promoted to operators, on each of the states in the Hilbert space,

$$\hat{\chi}_A \left|\Psi\right\rangle = 0. \tag{7.15}$$

However, first-class constraints often arise from gauge degrees of freedom, and it is frequently convenient to simplify the form of the equations of motion by choosing a specific gauge and maintaining that gauge choice throughout the computation. The strategy of gauge fixing is also supported by Dirac quantization. To fix the gauge, we introduce a complimentary set of constraints $\tilde{\chi}_A$, such that on the new set of constraints $\chi'_A = (\chi_A, \tilde{\chi}_A)$, the matrix of Poisson brackets

$$C_{AB} = \{\chi'_A, \chi'_B\}, \tag{7.16}$$

has nonzero determinant. At this point, we have a complete set of second-class constraints, and the techniques for second-class constraints, which are discussed below, can be used.

By contrast to the first-class case, second-class constraints cannot be directly quantized. Recall the toy example of constraints $q_1 = p_1 = 0$, which invalidate the desired commutation relation (7.13). In the case of second-class constraints, the Poisson bracket is a poor candidate for promotion to a commutator. Instead, the Dirac quantization program [402] develops a distinct bracket that is more well-behaved. Denote the nondegenerate matrix of the Poisson brackets of second-class constraints as

$$C_{AB} = \{\chi_A, \chi_B\}. \tag{7.17}$$

Note that this matrix may either be constructed from an expanded set of constraints in the case (7.16) of gauge-fixing, or arise directly from the structure of the theory itself. The *Dirac bracket* acting on a pair of functions F, G is defined as,

$$\{F, G\}_{\mathrm{DB}} = \{F, G\} - \{F, \chi_A\} C^{AB} \{\chi_B, G\}, \tag{7.18}$$

where C^{AB} is the matrix inverse of C_{AB},

$$C^{AB} C_{BC} = \delta^A{}_C. \tag{7.19}$$

It is easy to verify that the Dirac bracket (7.18) satisfies the desirable property that the constraints now have strongly vanishing Dirac brackets,

$$\{F, \chi_B\}_{\mathrm{DB}} = 0, \tag{7.20}$$

for all functions on phase space F. The theory may now be quantized by promoting the Dirac brackets to commutators, with the additional factor of $i\hbar$. Importantly, under the Dirac bracket, the constraints commute even off shell. This indicates, as detailed in [403], that there is no distinction between imposing the constraints on the physical states (7.15) and imposing the constraints at the operator level, so even after passing to the quantized theory, we may impose $\chi_A = 0$ throughout the computation.

To further illustrate the Dirac procedure, we consider a simple finite-dimensional system to work through the full set of steps. Consider a Lagrangian,

$$L(q_1, q_2, \dot{q}_1, \dot{q}_2) = \frac{1}{2}\dot{q}_1^2. \tag{7.21}$$

We construct a phase space as though the Lagrangian depended both on q_1 and q_2. The canonical momenta are then,

$$p_1 = \frac{\partial L}{\partial \dot{q}_1} = \dot{q}_1 \tag{7.22a}$$

$$p_2 = \frac{\partial L}{\partial \dot{q}_2} = 0, \tag{7.22b}$$

where we then identify the second equation as the primary constraint for the system $\chi_1 = p_2$.

We then construct the extended Hamiltonian,

$$\begin{aligned} H &= p_1\dot{q}_1 - \frac{1}{2}\dot{q}_1^2 + C_1 p_2, \\ &= \frac{1}{2}p_1^2 + p_2 C_1. \end{aligned} \tag{7.23}$$

Taking the Poisson bracket of the leading constraint with the Hamiltonian, we find that there are no further constraints to be found,

$$\{p_2, H\} = 0. \tag{7.24}$$

This constraint system is then of first-class: the single constraint trivially has vanishing Poisson bracket with itself. However, we know well that this system has a simpler description, so we adopt

the additional "gauge fixing" constraint $\chi_2 = q_2$. First, we verify that our new constraint condition does not give rise to secondary constraints,

$$\{q_2, H\} = C_1. \tag{7.25}$$

Instead of giving a new constraint, (7.25) simply sets to zero the free parameter C_1 of the Hamiltonian. We then have an adjusted Hamiltonian,

$$H = \frac{1}{2}p_1^2. \tag{7.26}$$

Our constraint system is now a second-class constraint system,

$$\{\chi_1, \chi_2\} = \{q_2, p_2\} = 1. \tag{7.27}$$

Therefore, we introduce the Dirac bracket associated with this constraint system,

$$\begin{aligned}\{F, G\}_{\mathrm{DB}} &= \{F, G\} - \{F, p_2\}\{q_2, G\} + \{F, q_2\}\{p_2, G\} \\ &= \frac{\partial F}{\partial q_1}\frac{\partial G}{\partial p_1} - \frac{\partial F}{\partial p_1}\frac{\partial G}{\partial q_1}.\end{aligned} \tag{7.28}$$

We see immediately that the Dirac bracket acts to remove the spurious degree of freedom from the original Lagrangian (7.21), leaving us with the Poisson bracket on the reduced space which we could have considered from the start. In the following sections, we construct a modification to the Dirac bracket which provides a different computational route to finding the dynamics of the physical phase space.

7.3 Modified Dirac bracket construction

7.3.1 Modified Dirac bracket construction for second-class constraint systems

In this section, we assume the constraint system χ_A is second-class. We define functions that are conjugate to χ_A as a set of functions ξ^A ("constraint conjugates") such that:

1.

$$\{\xi^B, \chi_A\} = \delta^B{}_A, \tag{7.29}$$

2. *For every function on phase space F which has vanishing Poisson bracket with the constraints χ_A on the surface $\chi_A = 0$:*

$$\{F, \chi_A\}\big|_{\chi_A=0} = 0, \tag{7.30}$$

that function F also has vanishing Poisson bracket with the conjugate constraints ξ^A on the constraint surface $\chi_A = 0$:

$$\{F, \xi^A\}\big|_{\chi_A=0} = 0. \tag{7.31}$$

We define the modified Dirac brackets using these conjugate constraints as

$$\{A, B\}_{\mathrm{mDB}} \equiv \{A, B\} - \{A, \xi^A\}\{\chi_A, \chi_B\}\{\xi^B, B\} - \{A, \chi_A\}\{\xi^A, B\} + \{A, \xi^A\}\{\chi_A, B\}. \tag{7.32}$$

For the following discussion, we use the induced 2-form on the constraint surface, so we review the salient details here. For any set of coordinates x^μ on phase space,

$$\{x^\mu, x^\nu\} = \sigma^{\mu\nu}. \tag{7.33}$$

On the full phase space, we can define an inverse,

$$\sigma_{\mu\nu}\sigma^{\nu\lambda} = \delta_\mu{}^\lambda. \tag{7.34}$$

We define a set of functions $x^\mu = x^\mu(y^i)$ which parametrically define the constraint surface, where y^i are coordinates on the constraint surface. The induced 2-form is obtained by deriving the pullback of the 2-form $\sigma_{\mu\nu}$ to the constraint surface,

$$\sigma_{ij} = \sigma_{\mu\nu}\frac{\partial x^\mu(y^i)}{\partial y^i}\frac{\partial x^\nu(y^j)}{\partial y^j}. \tag{7.35}$$

The induced 2-form σ_{ij} is covariant on indices y^i which parameterize the constraint surface and independent of the coordinates in which the functions x^μ are defined. In the special case in which only second class constraints are present, the induced 2-form may also be inverted,

$$\sigma_{ij}\sigma^{jk} = \delta_i{}^k, \tag{7.36}$$

which is the foundation of the definition of a preferred, coordinate-independent bracket on the constraint surface. The bracket associated with the induced 2-form is then

$$\{f(y^i), g(y^i)\}^* = \frac{\partial f(y^i)}{\partial y^i}\sigma^{ij}\frac{\partial g(y^i)}{\partial y^j}. \tag{7.37}$$

For the case of second-class constraints only, this is the bracket which we wish any modified bracket to reproduce. When first-class constraints are also present, the definition requires more refinement, as the induced 2-form σ_{ij} is then degenerate.

We review the argument for the original Dirac bracket reproducing the inverse of the induced 2-form in the special case of all second-class constraints in Section 7.3.1.1 and demonstrate that the modified Dirac bracket similarly reproduces the inverse of the induced 2-form when the system has only second-class constraints in Section 7.3.1.2. In Section 7.3.2, we use similar methods to demonstrate that the modified Dirac bracket reproduces the inverse of the induced 2-form on the reduced phase space in the first-class constraint case. We note that the proofs we present below demonstrate that the modified Dirac bracket and the Dirac bracket are identical on the constraint surface itself, but may differ off of the constraint surface. For second-class constraints, the modified Dirac bracket can therefore be constructed as the Dirac bracket defined on a nonlinear combination of the original constraints.

7.3.1.1 Dirac bracket validity for second-class constraints

The proof discussed in this section is simply a review of the proof given in [403] for the Dirac bracket, as a motivating warm-up for the similar proof given in Section 7.3.1.2 for the modified Dirac bracket. The Dirac bracket in this discussion is defined as

$$\{F, G\}_{\mathrm{DB}} = \{F, G\} - \{F, \chi_A\} C^{AB} \{\chi_B, G\}, \tag{7.38}$$

where

$$C^{AB}\{\chi_A, \chi_B\} = \delta^A{}_B. \tag{7.39}$$

Claim: *For second-class constraints, the restriction of the Dirac bracket to the constraint surface* $\chi_A = 0$ *is simply the induced bracket, that is*

$$\{F, G\}_{DB}\big|_{\chi_A=0} = \{F|_{\chi_A=0}, G|_{\chi_A=0}\}^*. \tag{7.40}$$

Proof: Both the right and left sides of the claimed equation (7.40) are coordinate-invariant, so we may demonstrate the equivalence in any particular convenient coordinates we like. We choose

the set of coordinates $x^\mu = (y^i, \chi_A)$ such that $\{y^i, \chi_A\}\big|_{\chi_A=0} = 0$, as may always be done [403]. In this case, the set of Poisson brackets is

$$\{x^\mu, x^\nu\} = \begin{pmatrix} \{y^i, y^j\} & \{y^i, \chi_B\} \\ \{\chi_A, y^j\} & \{\chi_A, \chi_B\} \end{pmatrix} = \begin{pmatrix} c^{ij} & 0 \\ 0 & C_{AB} \end{pmatrix}, \tag{7.41}$$

where each of the nonzero matrices in the final expression are nondegenerate. The symplectic two-form $\sigma_{\mu\nu}$ is the inverse of the full set of Poisson brackets,

$$\sigma_{\mu\nu} = \begin{pmatrix} \sigma_{ij} & 0 \\ 0 & C^{AB} \end{pmatrix}, \tag{7.42}$$

where

$$\sigma_{ij} c^{jk} = \delta_i{}^k, \tag{7.43a}$$

$$C^{AB} C_{BC} = \delta^A{}_C. \tag{7.43b}$$

In these coordinates, we also have the result that the Poisson brackets of $\{y^i, y^j\}$ are directly the inverse of the induced 2-form on the surface, $c^{ij} = \sigma^{ij}$.

Consider now the Dirac bracket for an arbitrary pair of functions F, G:

$$\begin{aligned} \{F, G\}_{\mathrm{DB}} &= \{F, G\} - \{F, \chi_A\} C^{AB} \{\chi_B, G\} \\ &= \frac{\partial F}{\partial y^i} \sigma^{ij} \frac{\partial G}{\partial y^j} + \frac{\partial F}{\partial \chi_A} C_{AB} \frac{\partial G}{\partial \chi_B} - \frac{\partial F}{\partial \chi_A} C_{AB} C^{BC} C_{CD} \frac{\partial G}{\partial \chi_D} \\ &= \frac{\partial F}{\partial y^i} \sigma^{ij} \frac{\partial G}{\partial y^j} = \{F|_{\chi_A=0}, G|_{\chi_A=0}\}^* \end{aligned} \tag{7.44}$$

7.3.1.2 Modified Dirac bracket validity for second-class constraints

Claim: *For second-class constraints, the restriction of the modified Dirac bracket to the constraint surface is the induced bracket:*

$$\{F, G\}_{mDB}\big|_{\chi_A=0} = \{F|_{\chi_A=0}, G|_{\chi_A=0}\}^* \tag{7.45}$$

Proof: First, we introduce a set of coordinates, as was used in the previous argument, $x^\mu = (y^i, \chi_A)$ such that the Poisson bracket of the full set of coordinates is,

$$\{x^\mu, x^\nu\} = \begin{pmatrix} \{y^i, y^j\} & \{y^i, \chi_B\} \\ \{\chi_A, y^j\} & \{\chi_A, \chi_B\} \end{pmatrix} = \begin{pmatrix} c^{ij} & 0 \\ 0 & C_{AB} \end{pmatrix} \tag{7.46}$$

In this block-diagonal form, it is immediate that the inverse of the induced 2-form is directly the upper-left block $\sigma^{ij} = c^{ij}$. Consider now the coordinate result of the condition 2 introduced above for second-class constraints. Due to the block-diagonal Poisson bracket, we have

$$\{f(y^i), \chi_A\}|_{\chi_A=0} = 0, \tag{7.47}$$

for any function of only the constraint surface coordinates $f(y^i)$. Therefore, condition 2 imposes

$$\{f(y^i), \xi^A\}|_{\chi_A=0} = 0. \tag{7.48}$$

Expanding the condition out in terms of coordinates, we obtain

$$\sigma^{ij} \frac{\partial \xi^A}{\partial y^j}\Big|_{\chi_A=0} = 0, \tag{7.49}$$

for all functions f, which implies $\partial \xi^A / \partial y^j|_{\chi_A=0} = 0$. Furthermore, condition 1 implies

$$\{\xi^A, \chi_B\}\big|_{\chi_A=0} = \frac{\partial \xi^A}{\partial \chi_C} C_{CB}\big|_{\chi_A=0} = \delta^A{}_B \tag{7.50}$$

Consider the Poisson bracket of an arbitrary function G with the conjugate constraint ξ^A,

$$\{\xi^A, G\}\big|_{\chi_A=0} = \frac{\partial \xi^A}{\partial \chi_C} C_{CB} \frac{\partial G}{\partial \chi_B}\Big|_{\chi_A=0} = \frac{\partial G}{\partial \chi_B}. \tag{7.51}$$

Finally, we are now ready to consider the modified Dirac bracket and compare it to the bracket associated with the induced 2-form $\{\}^*$. For arbitrary functions F and B, and in the carefully chosen coordinates, the modified Dirac bracket evaluates to

$$\begin{aligned} \{F, B\}_{\mathrm{mDB}}\big|_{\chi_A=0} &= \Bigg[\left(\frac{\partial F}{\partial y^i} \sigma^{ij} \frac{\partial G}{\partial y^j} + \frac{\partial F}{\partial \chi_A} C_{AB} \frac{\partial G}{\partial \chi_B} \right) + \left(\frac{\partial F}{\partial \chi_A} C_{AB} \frac{\partial G}{\partial \chi_B} \right) \\ &\qquad - \left(\frac{\partial F}{\partial \chi_A} C_{AB} \frac{\partial G}{\partial \chi_B} \right) - \left(\frac{\partial F}{\partial \chi_A} C_{AB} \frac{\partial G}{\partial \chi_B} \right) \Bigg] \Bigg|_{\chi_A=0} \\ &= \left(\frac{\partial F}{\partial y^i} \sigma^{ij} \frac{\partial G}{\partial y^j} \right) \Big|_{\chi_A=0} \\ &= \{F|_{\chi_A=0}, G|_{\chi_A=0}\}^*. \end{aligned} \tag{7.52}$$

To summarize the argument, the statement is that we can construct a set of coordinates in which the modified Dirac bracket simply acts to extract the component of the Poisson brackets which are directly identified with the induced 2-form bracket $\{\}^*$, and remove the rest. From the coordinate invariance of both objects we can conclude that they coincide in all coordinate systems:

$$\{F, G\}_{mDB}\big|_{\chi_A=0} = \{F|_{\chi_A=0}, G|_{\chi_A=0}\}^*, \tag{7.53}$$

as desired. Note that, as the modified Dirac bracket reproduces the same intrinsic quantity on the constraint surface, the two methods will produce identical results for second-class constraint systems. It is likely that computing modified Dirac brackets will be less involved for intricate systems, as it does not require inversion of a potentially large matrix of Poisson brackets. Additionally, modified Dirac brackets offer the possibility of simultaneously dealing with second-class and first-class constraints all at once, as we discuss below in Section 7.3.2.

7.3.2 Modified Dirac bracket construction for first-class constraints

In this section, we assume instead that the constraint system χ_A is first class. We define constraints conjugate to χ_A as a set of functions ξ^A such that:

$$\{\xi^B, \chi_A\} = \delta^B{}_A. \tag{7.54}$$

As in the case for second-class constraints, we define the modified Dirac bracket for a set of first-class constraints via the same formula (7.32),

$$\{F, G\}_{\text{mDB}} \equiv \{F, G\} - \{F, \xi^A\}\{\chi_A, \chi_B\}\{\xi^B, G\} - \{F, \chi_A\}\{\xi^A, G\} + \{F, \xi^A\}\{\chi_A, G\}. \tag{7.55}$$

In the case of first-class constraints, we must appropriately handle the gauge orbits on the constraint surface. It is these directions which cause the symplectic 2-form (7.35) to be non-invertible. Following [403], we introduce a set of coordinates z^α to parameterize a *reduced phase space*, such that the gauge orbits are all described by

$$z^\alpha(y^i) = \text{const.} \tag{7.56}$$

To emphasize the point, the value of the function z^α is independent of coordinate changes in directions along gauge orbits. In this way, the functions z^α are thought of as coordinates on the reduced phase space.

As shown in [403] Appendix 2A, the reduced phase space inherits an nondegenerate 2-form $\sigma_{\alpha\beta}$ from the induced 2-form σ_{ij}, which satisfies the property

$$\sigma_{ij} = \sigma_{\alpha\beta}\frac{\partial z^\alpha(y^i)}{\partial y^i}\frac{\partial z^\beta(y^i)}{\partial y^j}. \tag{7.57}$$

The nondegenerate 2-form $\sigma_{\alpha\beta}$ is then covariant in coordinates z^α and independent of the coordinate choice on the full constraint surface y^i. We may then define an inverse to the reduced 2-form $\sigma_{\alpha\beta}$,

$$\sigma_{\alpha\beta}\sigma^{\beta\gamma} = \delta_\alpha{}^\gamma \tag{7.58}$$

We then wish our modified bracket to reproduce the bracket associated with this inverse,

$$\{F, G\}^* = \frac{\partial F}{\partial z^\alpha}\sigma^{\alpha\beta}\frac{\partial G}{\partial z^\beta}, \tag{7.59}$$

which is independent of coordinates on phase space.

Claim: *For first-class constraints, the restriction of the modified Dirac bracket to the constraint surface $\chi_A = 0$ is the reduced bracket for gauge invariant functions. That is, for any pair of gauge-invariant functions F and G,*

$$\{F, G\}_{mDB}\big|_{\chi_A=0} = \{F, G\}^*\big|_{\chi_A=0} \tag{7.60}$$

where gauge-invariant functions satisfy

$$\{F, \chi_A\}\big|_{\chi_A=0} = 0 \tag{7.61}$$

Proof: For the case of first-class constraints, note that the requirement is actually far less stringent, as any pair of gauge-invariant functions F and G already has the property,

$$\{F, G\}|_{\chi_A=0} = \{F, G\}^*\big|_{\chi_A=0}. \tag{7.62}$$

We just need to verify that the definition of the modified Dirac bracket doesn't disrupt this desirable property. As we discuss below, the modified Dirac bracket also has other convenient properties in the case of first-class constraints which the Poisson bracket may not.

For gauge-invariant functions,

$$\{F, G\}\big|_{\text{mDB}} = \{F, G\} - \{F, \xi^A\}\{\chi_A, \chi_B\}\{\xi^B, B\}. \tag{7.63}$$

For first-class constraints,

$$\{\chi_A, \chi_B\}\big|_{\chi_A=0} = 0. \tag{7.64}$$

So, we have immediately that

$$\{F,G\}_{\mathrm{mDB}}\big|_{\chi_A=0} = \{F,G\}\big|_{\chi_A=0} - \left(\{F,\xi^A\}\{\chi_A,\chi_B\}\{\xi^B,B\}\right)\big|_{\chi_A=0} = \{F,G\}\big|_{\chi_A=0} = \{F,G\}^*\big|_{\chi_A=0}, \tag{7.65}$$

which is sufficient to prove the claim.

To close, we note that the modified Dirac bracket has the potentially desirable property that it ensures that all functions have strongly vanishing bracket with the constraint functions:

$$\{\chi_A, F\}_{\mathrm{mDB}} = \{\chi_A, F\} + \{\chi_A,\chi_B\}\{\xi^B, F\} - \{\chi_A,\chi_B\}\{\xi^B, F\} - \{\chi_A, F\} = 0. \tag{7.66}$$

In this way, the modified Dirac bracket provides a route to a commutation relation for which the constraints can be consistently imposed at operator level, regardless of whether the original constraints were first or second class. Finally, we also see that because an identical formula is used in both the first and second-class case, the modified Dirac bracket should deal gracefully with the case in which a set of constraints has small but nonvanishing brackets among one another.

The treatment should generalize straightforwardly to the case of mixed constraints. For mixed constraints, the second condition formulated in Section 7.3.1 must be applied only to the set of constraints which are second-class, as the additional condition leads to contradiction if applied to first-class constraints. We then expect that the corresponding Dirac bracket should reproduce the bracket associated with the reduced, induced 2-form on the surface for any gauge-invariant function on the surface.

7.4 | Conjugate constraints for perturbative Dirac states

The construction of conjugate constraints is also useful for the case in which all of the constraints χ_A are first-class,

$$[\chi_A, \chi_B] \approx 0. \tag{7.67}$$

In this case, the Dirac quantization program suggests the two options discussed in Section 7.2.2. One option is to introduce an equal number of additional constraints $\tilde{\chi}^A$, which together with the

original constraints form a system of second-class constraints $\chi'_A = (\chi_A, \tilde{\chi}_A)$. The alternative is to construct a state which simultaneously satisfies all of the original constraints,

$$\chi^A |\Psi\rangle = 0, \tag{7.68}$$

provided the constraints weakly commute, permitting the constraint condition (7.68) to be well-defined.

For perturbative expansions of the action, as are found often in effective field theories, it is valuable to develop a simple order-by-order procedure such that, given a perturbative expansion of the constraints,

$$\chi^A = \chi^{(0)A} + \varepsilon\chi^{(1)A} + \varepsilon^2\chi^{(2)A} + \mathcal{O}(\varepsilon^3), \tag{7.69}$$

in powers of a small parameter ε, and a 'free state' $\left|\Psi^{(0)}\right\rangle$,

$$\chi^{(0)A} \left|\Psi^{(0)}\right\rangle = 0, \tag{7.70}$$

to find the order-by-order expansion of the 'dressed state',

$$|\Psi\rangle = \left|\Psi^{(0)}\right\rangle + \varepsilon \left|\Psi^{(1)}\right\rangle + \varepsilon^2 \left|\Psi^{(2)}\right\rangle + \mathcal{O}(\varepsilon^3), \tag{7.71}$$

such that $|\Psi\rangle$ satisfies the full constraints (7.68). To find an appropriate set of dressed states, we will require the perturbative form of the weak commutation of constraints (7.67),

$$[\chi^{(0)A}, \chi^{(0)B}] \approx 0, \tag{7.72a}$$

$$[\chi^{(1)A}, \chi^{(0)B}] + [\chi^{(0)A}, \chi^{(1)B}] \approx 0, \tag{7.72b}$$

$$[\chi^{(2)A}, \chi^{(0)B}] + [\chi^{(1)A}, \chi^{(1)B}] + [\chi^{(0)A}, \chi^{(2)B}] \approx 0, \tag{7.72c}$$

as well as several other identities which follow from these and the Jacobi identity,

$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0. \tag{7.73}$$

We iteratively construct the dressed state using the commutators of the perturbative constraints and the conjugates to the leading order constraints ξ^A. The first order dressed state is then found as,

$$\left|\Psi^{(1)}\right\rangle = \xi^A\chi^{(1)}_A \left|\Psi^{(0)}\right\rangle + \frac{1}{2}\xi^A\xi^B \left[\chi^{(0)}_A, \chi^{(1)}_B\right] \left|\Psi^{(0)}\right\rangle + \frac{1}{6}\xi^A\xi^B\xi^C \left[\chi^{(0)}_A, \left[\chi^{(0)}_B, \chi^{(1)}_C\right]\right] \left|\Psi^{(0)}\right\rangle \tag{7.74}$$

where the indices A are raised and lowered with δ_{AB} and Einstein summation convention is used. It is straightforward to verify that this satisfies the desired first order constraint equations,

$$\chi^{(0)A}\left|\Psi^{(1)}\right\rangle = -\chi^{(1)A}\left|\Psi^{(0)}\right\rangle. \tag{7.75}$$

Next, we perform the iterative expansion to second order, yielding the dressed state,

$$\begin{aligned}
\left|\Psi^{(2)}\right\rangle = {} & \xi^A\chi_A^{(2)}\left|\Psi^{(0)}\right\rangle + \frac{1}{2}\xi^A\xi^B\left[\chi_A^{(0)},\chi_B^{(2)}\right]\left|\Psi^{(0)}\right\rangle + \frac{1}{6}\xi^A\xi^B\xi^C\left[\chi_A^{(0)},\left[\chi_B^{(0)},\chi_C^{(2)}\right]\right]\left|\Psi^{(0)}\right\rangle \\
& + \frac{1}{24}\xi^A\xi^B\xi^C\xi^D\left[\chi_A^{(0)},\left[\chi_B^{(0)},\left[\chi_C^{(0)},\chi_D^{(2)}\right]\right]\right]\left|\Psi^{(0)}\right\rangle + \xi^A\chi_A^{(1)}\left|\Psi^{(1)}\right\rangle \\
& + \frac{1}{2}\xi^A\xi^B\left[\chi_A^{(0)},\chi_B^{(1)}\right]\left|\Psi^{(1)}\right\rangle + \frac{1}{6}\xi^A\xi^B\xi^C\left[\chi_A^{(0)},\left[\chi_B^{(0)},\chi_C^{(1)}\right]\right]\left|\Psi^{(1)}\right\rangle \\
& + \frac{1}{2}\xi^A\xi^B\chi_A^{(1)}\chi_B^{(1)}\left|\Psi^{(0)}\right\rangle + \frac{1}{2}\xi^A\xi^B\xi^C[\chi_A^{(0)},\chi_B^{(1)}]\chi_C^{(1)}\left|\Psi^{(0)}\right\rangle \\
& + \frac{1}{6}\xi^A\xi^B\xi^C[\chi_A^{(1)},[\chi_B^{(0)},\chi_C^{(1)}]]\left|\Psi^{(0)}\right\rangle + \frac{1}{6}\xi^A\xi^B\xi^C\xi^D[\chi_A^{(0)},[\chi_B^{(0)},\chi_C^{(1)}]]\chi_D^{(1)}\left|\Psi^{(0)}\right\rangle \\
& + \frac{1}{8}\xi^A\xi^B\xi^C\xi^D[\chi_A^{(0)},\chi_B^{(1)}][\chi_C^{(0)},\chi_D^{(1)}]\left|\Psi^{(0)}\right\rangle + \frac{1}{24}\xi^A\xi^B\xi^C\xi^D[\chi_A^{(0)},[\chi_B^{(1)},[\chi_C^{(0)},\chi_D^{(1)}]]\left|\Psi^{(0)}\right\rangle \\
& + \frac{1}{12}\xi^A\xi^B\xi^C\xi^D\xi^E[\chi_A^{(0)},[\chi_B^{(0)},\chi_C^{(1)}]][\chi_D^{(0)},\chi_E^{(1)}]\left|\Psi^{(0)}\right\rangle \\
& + \frac{1}{72}\xi^A\xi^B\xi^C\xi^D\xi^E\xi^F[\chi_A^{(0)},[\chi_B^{(0)},\chi_C^{(1)}]][\chi_D^{(0)},[\chi_E^{(0)},\chi_F^{(1)}]]\left|\Psi^{(0)}\right\rangle.
\end{aligned} \tag{7.76}$$

We conjecture that this iterative procedure can be pushed to arbitrarily high order, permitting a perturbative construction of the states $\left|\Psi^{(n)}\right\rangle$ dressed by the coupling between the source and gauge fields, giving increasing powers of various combinations of $\xi^A\chi_A^{(n)}$.

7.5 | Conclusions

In this chapter, we have presented a novel set of techniques for approaching the methods of Dirac quantization. Using the construction of a set of conjugate constraint functions, we find both an iterative method for assembling a constraint-satisfying physical state, and a method for defining a modified formulation of the Dirac bracket. Provided the existence of a set of conjugate constraints, our modified Dirac bracket formalism offers a more direct bracket construction which nonetheless permits quantization.

Chapter 7 Bibliography

[402] P. A. M. Dirac, *Lecturees on Quantum Mechanics* (Dover, 1931)

[403] M. Henneaux and C. Teitelboim, *Quantization of gauge systems* (1992)

[404] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory* (Addison-Wesley, Reading, USA, 1995),
URL: http://www.slac.stanford.edu/~mpeskin/QFT.html

[405] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations* (Cambridge University Press, 2005)

[406] J. Lee and R. M. Wald, *Local symmetries and constraints*, J. Math. Phys. **31** (1990), pp. 725–743

[407] J. D. Reyes, *Canonical methods in classical and quantum gravity: An invitation to canonical LQG*, J. Phys. Conf. Ser. **1010** (2018) (1), p. 012001

[408] A. Ashtekar and J. Pullin (eds.), *Loop Quantum Gravity*, vol. 4 of *100 Years of General Relativity* (World Scientific, 2017)

[409] J. Hallin, *QED in (1+1)-dimensions by Dirac quantization* (1993)

[410] N. S. Baaklini and M. Tuite, *Dirac Quantization of Spin-2 Field*, J. Phys. **A12** (1979), p. L13

[411] N. S. Baaklini and M. Tuite, *Dirac Quantization of Massive Spin 3/2 Field*, J. Phys. **A11** (1978), p. L139

[412] T. Darkhosh, *Dirac quantization of massive spin one particles in an extremal symmetrical tensor field*, J. Math. Phys. **31** (1990), pp. 2201–2203

[413] F. Ardalan, H. Arfaei and M. M. Sheikh-Jabbari, *Dirac quantization of open strings and noncommutativity in branes*, Nucl. Phys. **B576** (2000), pp. 578–596

[414] M. Varadarajan, *Dirac quantization of parametrized field theory*, Phys. Rev. **D75** (2007), p. 044018